UNIVERSITY OF NAIROBI

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

SCHOOL OF MATHEMATICS

RESEARCH PROJECT IN MSC BIOMETRY

COMPARISON OF HIV/AIDS MORTALITY RATES IN AN URBAN AND A RURAL SETTING.

BY

JOAN WAKARINDI THIGA

REG NUMBER: I586/70537/2007

A PROJECT SUBMITTED TO THE SCHOOL OF MATHEMATICS IN PARTIAL FULFILLMENT FOR A DEGREE OF MASTER OF SCIENCE IN BIOMETRY

JULY 2009

Declaration

I the undersigned declare that this project is my original work and has not been presented as a degree in any other university

Joan Wakarindi Thiga

Reg No: I586 / 70537 / 07

Signature ..... Date 11/08/09

The project has been submitted for examination with my approval as the university supervisor

Prof Manene,

School of mathematics,

University of Nairobi,

P.O Box 30197,

Nairobi (kenya)

Signature ..... Date 13/08/09

Prof J.A. M. Otieno,

School of mathematics,

University of Nairobi,

P.O Box 30197,

Nairobi (Kenya)

Signature ..... Date. 14/08/2009.....

Dedication

To my husband Naftali, son Brian and parents Mr. & Mrs. Thiga.

Acknowledgement

I thank the almighty God for the good health and a sound mind thought my research work.

I greatly appreciate the guidance from my Supervisors Prof Manene and Prof J.A.M Atieno. It is through your support that I was able to learn and complete my research successfully.

To my classmates in the MSC Biometry class for the continuous encouragement through the entire course. You were a great team.

To my husband, Parents and siblings for your moral support.

Table of Contents

Chapter one

1.1 Introduction

HIV/AIDS:

HIV (human Immuno deficiency Virus) and AIDS (Acquired Immune Deficiency Syndrome) is the most devastating disease that has ever faced the human kind. In 2007, it was estimated that 33.2 million people lived with HIV virus and it had killed 2.1 million adults and 330,000 children. Three quarters of these deaths occurred in sub-Saharan Africa (UNAIDS, WHO,2007). New infections are increasing everyday especially among the low income countries and among the most vulnerable groups in these countries.

The HIV virus which is transmitted through blood or other body fluids destroys the white blood cells which protect the body against diseases. The body becomes weak and prone to diseases such as tuberculosis, meningitis among other life threatening diseases. AIDS eventually creeps in which is evidenced by the presence of many diseases in the human's body due to the weak body immunity.

The indicators used to measure the disease progression in one's body are CD4 count, viral load and the World Health Organization (WHO) stage. When the CD4 count is low or the viral load too high one is initiated on antiretroviral therapy (ARVS). ARVS are a combination of certain drugs which are given to the patient to slow down the rate at which the virus destroys the white blood cells. The body is then in a position to fight the

opportunistic infections. If well adhered to, the ARVS help the patient to lead a normal life.

HIV/AIDS in Kenya:

Kenya is experiencing a devastating HIV/AIDS epidemic, high maternal and infant morbidity and mortality, and deteriorating financial and human resources for health services. The HIV/AIDS is a threat to Kenya's development.

About two thirds of the Kenyan population lives below the poverty line. They depend on health care services provided by government funded health institutions which are poorly equipped. This is unlike the wealthy who can afford treatment in private health institutions which are better equipped. The intricate relationship between poverty line and HIV continues to be a vicious cycle in the national response to the pandemic.

The increase in the spread of the virus can be attributed to the increase in poverty levels. In the quest to earn a living, the poor become vulnerable since they engage in risky behaviors such as prostitution, sharing of needles, sharing of razor , and injections needles is also common among the poor.

According to the ministry of health Kenya records, urban areas in Kenya reported higher prevalence levels than rural areas in 1995 but the situation reversed between 1999 and 2002. The reversal in prevalence rates was due to PLWHA (people living with HIV/AIDS) returning to the rural areas to seek support from relatives when they get laid off as result of their decline in productivity. Urban dwellers have more access to health

care institutions for medical supplies and are more likely to keep abreast with new research findings as compared to the rural dwellers

The socio economic impact of HIV/AIDS has manifested itself in key areas of the economy. There is evidence that socio economic variables are correlated with life expectancy. Health expenditures have risen and will continue to rise and productivity will be adversely affected .Controlling HIV/AIDS epidemic is a key part of strengthening the economy in the 21st century.

The government should reshape policy to redistribute resources objectively. Since HIV/AIDS affects all sectors, a multidisplinary approach should be initiated in planning and budgeting for HIV management. HIV/AIDS should be reviewed as a developmental problem and therefore tackled by all arms of the government over and above the effects of other stakeholders. In the recent years, countries, foundations and developed institutions have donated funds to finance anti AIDS projects in Kenya.

1.2 Problem Statement

HIV/AIDS affects all sectors of the economy and developing countries recorded heavy losses due to sickness and death. Despite the fact that a high % of the population is already infected, the fight still revolves around prospective and preventive measures. There is lack of agility to match the speed of spread and effects of HIV especially in poor communities. Based on existing information about survival rates, developed countries

like U.S.A have sustained PLWHA for 10-15yrs compared to their underdeveloped African counterparts 2-3yrs.

HIV/AIDS affects mortality rates which results in lower national life expectancy. It causes higher infant mortality, decreased productivity of the working population, and changes in the distribution of population by age and sex *(Family life Counseling Association Of Kenya, 1996).* Prevalence rates have been shown to have some trends as selected socio-economic factors like poverty, female literacy rate, gross domestic product and the geographical variations exist.

The Kenya Central Bureau of Statistics estimates that without AIDS, life expectancy at birth would currently be about 65 years. However, because of the large number of AIDS deaths, life expectancy at birth is actually only about 46 years and may decline to 45 years by 2010. Thus almost 20 years of life expectancy have already been lost because of *AIDS. (AIDS in Kenya, Kenya Ministry of Health, 2001 sentinel surveillance system).*

In sub-Saharan Africa, estimated life expectancy is 47yrs down by 5yrs since 1993 and an estimated 15yrs shorter than it would have been in the absence of AIDS (UNAIDS/WHO, 2000). There is fear that life expectancy may decline further if poverty levels and the HIV/AIDS infection rate continue increasing.

1.3 Hypothesis:

Null hypothesis: Patients infected with HIV/AIDS in the rural setting have the same mortality rate than those in an urban setting.

Alternative hypothesis: Patients infected with HIV/AIDS in the rural setting have a higher mortality rate than those in an urban setting.

1.4 Objectives:

a. To determine whether the mortality rate between the patients in the rural setting and those in an urban setting are the same.

b. To determine the main factors associated with mortality rates.

c. To determine whether the main factors associated with mortality rates are the same in the two settings.

1.5 Literature review

The emergence of AIDS has led to renewed interest in adult mortality especially given that adults aged between 15 – 60 form the reproductive and productive group and are responsible for the welfare of the younger and older groups. This group represents more than 50% of the population of the sub-Saharan Africa (United nations 2006).

The life of a HIV infected person can be prolonged by taking of ARVS. The ARVS boosts the person's immunity. Several counseling sessions are take before and after

taking the drugs. They are aimed at ensuring the patient takes medications as prescribed. This process is known as the HAART process. The body therefore can fight away some life threatening diseases which attacks the body leading to death.

Several factors could limit the effectiveness of Highly Active anti-retroviral Therapy (HAART) in resource poor settings. Interruptions in supply at the program level or patients' limited financial resources might compromise adherence and treatment efficacy. Lack of proper diet which is key for the ARVS to be effective, lack of knowledge on how best to take care of oneself among others. The high prevalence of co – infections, notably tuberculosis and other bacterial diseases might also affect prognosis.

According to center for disease control and prevention, 1993, rural area aids cases are increasing at approximately 3 times the rate in the urban areas. This could be attributed to the HIV infected moving to the rural areas from the urban centers. With the rise in number of the infected and the limited resources in the rural area, it is interesting to study the patient's progress on being initiated on ARV.

Previous studies identified problems confronting HIV infected residents as long distances to the medical facilities, increased discrimination and cost/ accessibility of life extending drugs. The studies indicated that compared with HIV infected urban residents, people with HIV in the rural areas would report lower quality of life, more barriers to care, elevated perception of loneliness and more frequent incidences of AIDS related infections.

A retrospective study of prospectively collected data from consecutively enrolled adult patients in nine HIV clinics in western Kenya was done. The study aimed at determining the predictors of mortality among the HIV infected patients. Data from records of deceased patients started on HAART between November 2001 and December 2005 were analyzed and compared with those from records of living patients started on HAART during the same period. A comparison was made between 527 deceased patients and 1054 alive patients. The study's results were, Median age was 38yrs (range 16 -77 years) for the deceased and 36years (range 15 -73years) for the alive patients. Median duration for the time on HAART was 7.7weeks (range 0 - 110) and 42 weeks (range 0-28) ($p<0.005$) for the deceased and the alive patients respectively. Patients with CD4 count <100 were more likely to die than those with CD4 count 100-200 (HR =1.94, CI (1.63 ,2.53) $p <0.001$). The hazard for death of perfectly adherent patients (HR =0.61, CI (0.44, 0.86) , $p =0.0025$) while that for patients attending urban clinic was a third of that among the patients attending a rural clinic.

F dabis and Megger(March,2006) carried out a comparative study between the low income setting (Africa, Asia and south America) and high income settings (Europe, North America) HiV-1 mortality rates. The study was done in the first year of antiretroviral therapy (ARV) in the HIV infected patients. They studied 4810 treatment naïve adults patients from the low income setting and 22,217 patients from the high income setting. The study's results were, Lower CD4 cell count was noted among the low income setting (median 108 cells per μl) versus (median 234 cells per μl) among the high income

setting. Six months later, the median number of cells gained was (106 cells per μl Vs 103 cell per μl). Mortality was higher in low income setting than in high income setting. Provision of free treatment in low income setting was associated with lower mortality (adjusted HR 0.23; 95% CI 0.80-0.61). The study's conclusion was that patients starting HAART in low income setting have high mortality rates in the first months of therapy as compared to patients in high income settings. Timely diagnosis of the disease , assessment of treatment eligibility and provision of free HAART might reduce the mortality rate.

In England, a retrospective cohort study involving 644 HIV infected adults was done. The study's aim was to explain the increased mortality rates in patients infected with HIV living in the rural areas. The study compared mortality rates in 327 patients living in the rural areas with 317 patients living in the urban areas. A multivariate logistic regression model was used. It was found out that patients in the rural areas with HIV infection were older at the end of the follow up (43.4 Vs 41.4 years, p =0.002), the mean CD4 count at presentation was similar in the two groups (376 Vs 351 cells / μl p= 0.298). Mortality was higher in rural patients. The risk remained higher in the rural patients when adjusting for age, sex, race, HIV risk factors, year of diagnosis, receipt of antiretroviral treatment in a logistic regression model (OR 2.11, 1.064 to 4.218, p =0.047). The stud's conclusion was that HIV patients living in the rural areas have higher mortality rate than those living in the urban areas.

1.6 Significance of the Study

Understanding attributes to mortality can help in formulating strategies and establishing priorities to improve care. The following groups will benefit from the study:

Health workers

The results of the study will help the health workers who deal with the HIV/AIDS patients identify which factors are associated with mortality rates. This will enable them reduce the deaths caused by factors that can be controlled through educating their patients, reading widely for any new discoveries .e.t.c.

Government

From the study results, the government can strategize and campaign against the factors associated with mortality. The government can also offer support through funding projects aimed at addressing these factors.

The infected and affected persons

The results will create awareness to the infected people and the affected people on factors which are mainly associated with mortality rates. They will in turn be in a position to take better care of themselves. The affected ones will know how best to take care of their loved ones thus reducing the mortality rate.

Chapter Two

2.1 Methodology:

A retrospective study on a group of patients enrolled in a rural and an urban clinic between Jan 2006 up to Dec 2008 will be used. The Kaplan Meier, log rank test and the Cox PH model were used in the analysis. The patients included in the analysis had met the criteria below:

- The patient was ARV naïve on enrollment.
- On enrollment to the clinic, the patient was eligible for ARV
- The patient was subsequently initiated on ARV.
- The patient was 16yrs and above.

2.2 The Kaplan Meier

Let s(t) be the probability that an item from a given population will have a lifetime exceeding t. For a sample of size n from this population, let the observed times until death of n sample members be

$$t_1 \leq t_2 \leq t_3 \ldots\ldots\ldots \leq t_n$$

Corresponding to each t_i is :

n_i, the number "at risk" just prior to time t_i and

d_i the number of deaths at time t_i

The intervals between each time typically will not be uniform.

The Kaplan Meier is a product of the form

$$s(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i}$$

Kaplan Meier estimates the survival function from life time data. It is the non- parametric maximum likelihood estimator of s(t). It can be used to measure the fraction of patients living for a certain amount of time after treatment.

In case of no censoring, n_i is just the number of survivors just prior to time t_i while with censoring, n_i is the number of survivors less the number of censored cases.

$t_1, t_2, t_3 \ldots\ldots\ldots, t_n$ are the different time points at which death occurs for the HIV patients.

d_i is the number of deaths that occur at time point t_i.

n_i is the survivors at the time point t_i .

Only the surviving cases are observed. Kaplan meier curve is based on the following assumptions:

- Censoring is unrelated to prognosis.
- Survival probabilities are the same for subjects recruited early and those recruited late.
- Events happened at the specified times.

Comparison between the two Kaplan Meier curves will be done using the log rank test.

2.3 Log rank test.

Log rank test is a hypothesis test to compare the survival distributions of two or more samples. It is a non – parametric test whose null hypothesis is that the risk of death is the same in all groups. It is a test of significance and cannot provide an estimate of the size of the difference between the groups.

The log rank test statistic compares estimates of the hazard functions of the two groups at each observed event time. It is constructed by computing the observed and the expected numbers of event time and then adding this to obtain an overall summary across all the time points where there is an event.

Let $j = 1, 2 \ldots n$ be the distinct times of observed events in either group for each time j.

Let n_{1j} and n_{2j} be the number of subjects "at risk" at the start of period j in the groups respectively.

Let O_{1j} and O_{2j} be the observed number of events in the groups respectively at time j and define

$O_j = O_{1j} + O_{2j}$

Given that O_j events happened across both groups at time j under the null hypothesis, O_{1j} has the hyper geometric distribution with parameters n_j , n_{1j} and O_j The distribution has expected value:

$$E_j = O_j \frac{n_{1j}}{n_j}$$

Variance:

$$V_j = \frac{O_j(n_{1j}/n_j)(1-n_{1j}/n_j)(n_j-O_j)}{n_j-1}$$

The log rank statistic compares each O_{1j} to it's expectation E_j under the null hypothesis and is defined as:

$$z = \frac{\sum_{j=1}^{J}(O_{1j} - E_j)}{\sum_{j=1}^{J} V_j}$$

2.4 Cox proportional hazard models

Often one is interested in comparing two or more groups of time to time event. Subjects may have demographic variables recorded such as age, gender, socio-economic status, education variables, physiological variables such as blood pressure, glucose level .e.t.c. Such variables may be used as covariates (explanatory variables, confounders, risk factors, independent variables) explaining the response (dependent) variable. After adjustment for those potential explanatory variables, the comparison of survival times between groups should be less biased..

The multiplicative hazards model due to cox (1972) often referred to as "proportional hazards model" is commonly used. The Cox proportional hazard models were introduced by Dr Cox in order to estimate the effects of differential covariates influencing the time to failure of a system. It has been widely used in the biomedical field. It is a non parametric model. It is not based on any assumptions concerning the nature or shape of the underlying survival distribution.

Let X denote the time to some event. The data based on a sample of size n, consists of the triple $(T_j, \delta_j, Z_{j(}(t)); j = 1....n)$ where:

T_j = Time on the study for the patient

δ_j = Event indicator for the j^{th} patient . $\delta_j = 1$ if the event has

Occurred, $\delta_j = 0$ if the time is right censored.

$Z_t = (Z_{j1}(t),\ldots\ldots, Z_{jp}(t))^t$ is the vector of covariates or the risk

Factors for the j^{th} individual at the time t which may affect the

survival distribution of X. The covariates may be time dependent

or not.

Let h(t/**z**) be the hazard rate at time t for an individual with risk vector **Z.** Then the basic model due to Cox is:

$$h(t / z) = h_o(t)c(\beta' z)$$

Where h_o : an arbitrary hazard rate.

$\beta = (\beta_1 \ldots\ldots \beta_p)^t$ is a known parameter vector

$c(\beta' z)$ is a known parameter function

The model is a semi – parametric model since a parametric model form is assumed only for the covariate effect. The baseline hazard rate is treated non-parametrically since $h(t / z)$ must be positive.

A common model for $c(\beta' z)$ is

$$c(\beta' z) = exp(\beta' z) = exp(\sum_{k=1}^{p} \beta_k z)$$

Yielding

$$h(t / z) = h_o(t)\exp(\beta' z) = h_o(t)\exp(\sum_{k=1}^{p} \beta_k z_k)$$

The Cox model is often called a proportional hazards model because two individuals with covariate values z and $\mathbf{z}^*$, the ratio of their hazard rate is

$$\frac{h(t/z)}{h(t/z^*)} = \frac{h_0(t)\exp(\sum_{k=1}^{p}(\beta_k z_k))}{h_0(t)\exp(\sum_{k=1}^{p}(\beta_k z^*_k))} = \exp\sum_{k=1}^{p}\beta_k(z_k - z^*_k) \ldots\ldots\ldots(1)$$

Which is a constant, So the hazard rates are proportional.

The quantity (1) is called the relative risk (hazard ratio) of an individual with risk factor **Z** having the event as compared to an individual with risk factor $\mathbf{Z}^*$.

The goal is to make an inference about β in a global sense. The inferences are made based on whether there are distinct event times in the data or ties are present in the data. They are based on a partial or conditional likelihood rather than a full likelihood approach.

2.4.1 Partial likelihood when ties are present

Let $t_1 < t_2 < \ldots\ldots\ldots\ldots t_D$ denote D distinct, ordered event times. Further let d_i be the number of deaths at t_i and $\mathbf{D}_i$ the set of all individuals who die at time t_i. Let $\mathbf{S}_i$ be the sum of the vectors $\mathbf{Z}_j$ over all individual who die at t_i.

$$S_i = \sum_{j \in D_i} z_j$$

Let $\mathbf{R}_i$ be the set of all individuals at risk just prior to t_i. The partial likelihood due to cox(1972) is based on a discrete time, hazard rate model. This likelihood is constructed by assuming a logistic model for the hazard rate.

If we let h(t/**z**) be the conditional death probability in the interval (t,t+1) given survival to the start of the interval and we assume

$$\frac{h(t/z)}{1-h(t/z)} = \frac{h_0(t)}{1-h_0(t)}\exp(\beta' z)$$

Then this is the proper partial likelihood. Let Q_i denote the set of all subjects of d_i individuals who could be selected from the risk set R_i Each element of Q_i is a di –tuple of individuals who could have been one of the d_i failures at time t_i . Let $q = (q_i \ldots\ldots q_{di})$ be one of the elements of Q_i and define

$$s_{q^*} = \sum_{j=1}^{dj} z_{qj}$$

Then the discrete log likelihood is given by

$$l_3(\beta) = \prod_{i=1}^{D} \frac{\exp(\beta' s_i)}{\sum_{q \in Qi} \exp(\beta' s^*_q)}$$

2.4.2 Model building using the proportional hazard model

Two distinctly different important problems in regression are:

- Adjusting for potential confounders (or explanatory) variables when one has a specific hypothesis and the desire is to compare two or more groups with respect to survival times.

- To predict the distribution of the time to some event from a list of explanatory variables with no particular hypothesis.

When one has a particular hypothesis

The model building is aimed at adjusting that particular comparison(s) for the uncontrollable factors. The other explanatory factors are viewed as adjusters or confounders and interest in them matters only as far as they affect the assessment of the basic hypothesis.

Perform the global test of the primary hypothesis. It gives an impression of the simple, unadjusted relationship between the basic hypothesis factor and survival.

Consider the relationship between each of the other explanatory factors and survival, given that the factor stated in the basic hypothesis is already in the model.

When one has no particular hypothesis in mind

The aim is to obtain a set of variables which will aid in modeling survival or identifying a set of variables which may be used in testing a hypothesis (hypothesis generating)

Perform separate global tests for each explanatory factor so as to examine

the simple relationship between the explanatory variable and survival. This helps to determine which factor is most related to survival.

Consider the relationship between each of the other explanatory factors (not the one identified the most significant one) and survival given that the factor identified as the most significant is already in the model.

Chapter Three

3.1 Application of the model and the results

The outcome variable is mortality of the HIV patients under treatment. The key predictor variables are:

Type of clinic: The type of clinic was determined by the geographical location of the clinic. It could be either an urban or a rural clinic.

Age: This is the age of the patient involved in the study.

Gender: This is the gender of the patients involved in the study.

Income – This is the patient's monthly income. It is given in three categories namely:

- < Ksh5,000
- Ksh5 ,001 – 10,000
- > Ksh10,000

Education – This is the patient's highest level of education. This is given in three in three different categories:

- Lower than five years of primary education
- Five to eight years of primary education
- Beyond primary education

Disclosure – This helps to indentify patients who have disclosed their status to at leat one person. It is a categorical variable of either Yes or No.

WHO stage on enrollment – This is the stage in which the disease is at, at the time of seeking care in to the clinic. It is a categorical variable with stage 1&2 and stage 3&4.

Baseline Weight – This is the patient's weight at the time of seeking care at the clinic. It is in kg.

Baseline CD4 – This is the patient's cd4 count at the time of seeking care at the clinic. It is measured in cells/μl

Adherence – This is a measure of how well the patient takes the medications. It is a categorical variable:

- Ever missed more than half the doses
- Never missed more than half the doses

Month 6&12 CD4 – This is the CD4 at the 6th & the 12th months after taking the ARVS.

Month 6&12 weight – This is the weight at the 6th & the 12th months after taking ARVS.

3.2 Fitting the data to the Kaplan Meier function

The Kaplan Meier was used to estimate the survival function. The main objective of the Kaplan Meier is to determine whether the mortality rates are the same in the two settings under study. The function has the form

$$s(t) = \prod_{t_i \leq t} \frac{n_i - d_i}{n_i},$$

t_i = Time point at which death occurred.

n_i = Number of patients at risk at time point t_i (The number of patients who have not yet died or censored at time point t_i)

d_i = Number of deaths at time point t_i

3.3 Fitting the data to the Cox proportional hazard model

The model was used to determine which factors are associated with mortality. The model has the form

$$h(t/z) = h_o(t)\exp(\beta' z) = h_o(t)\exp(\sum_{k=1}^{p} \beta_k z_k)$$

$h_o(t)$ = It's an arbitrary hazard rate

z_k = is the vector of covariates (age, gender, disclosure, education, income, weigh, cd4 and the WHO stage)

β_k = it is the vector to be determined and inference made on

$\exp(\beta)$ Is the HR (hazard ratio) of the risk of the event for an individual in the rural setting relative to that of an individual in an urban setting for categorical variables e.g. income. For quantitative variables, e.g. weight, it is the risk of death for an individual with 'x'kg relative to an individual with 'x+1'kg

Chapter Four

4.1 Data analysis and Interpretation

Chart 1: A histogram for the baseline weight

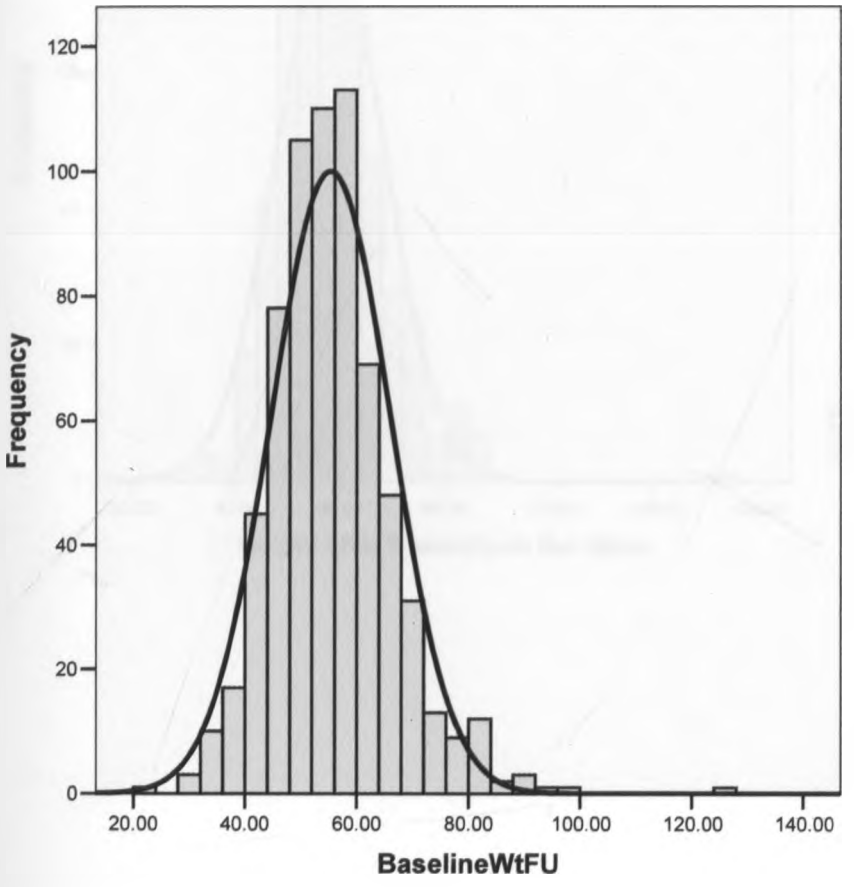


Chart 2: A histogram for weight after six months in the clinic

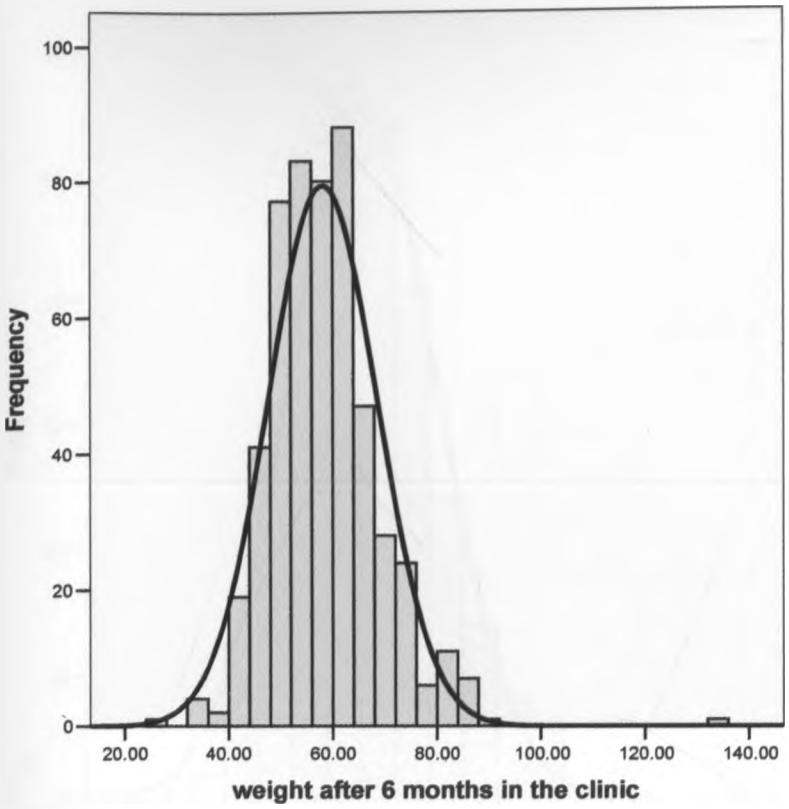


Chart 3: A histogram for weight after twelve months in the clinic

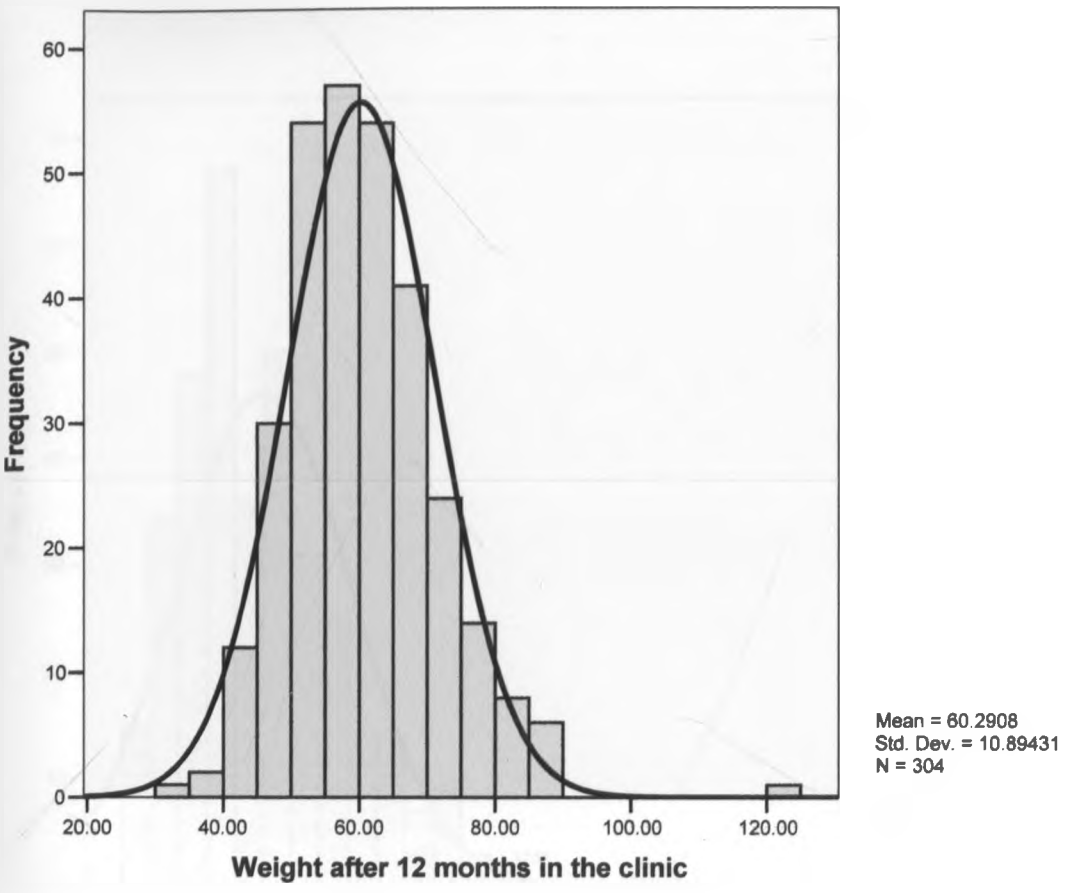


Chart 4: A histogram for cd4 after six months in the clinic

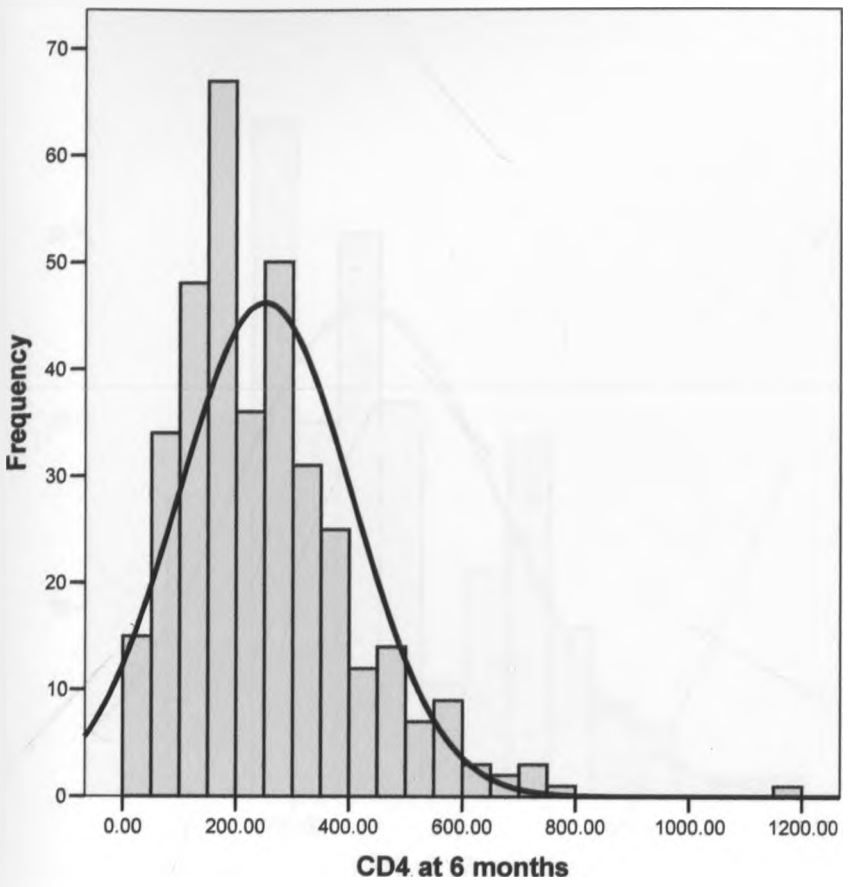
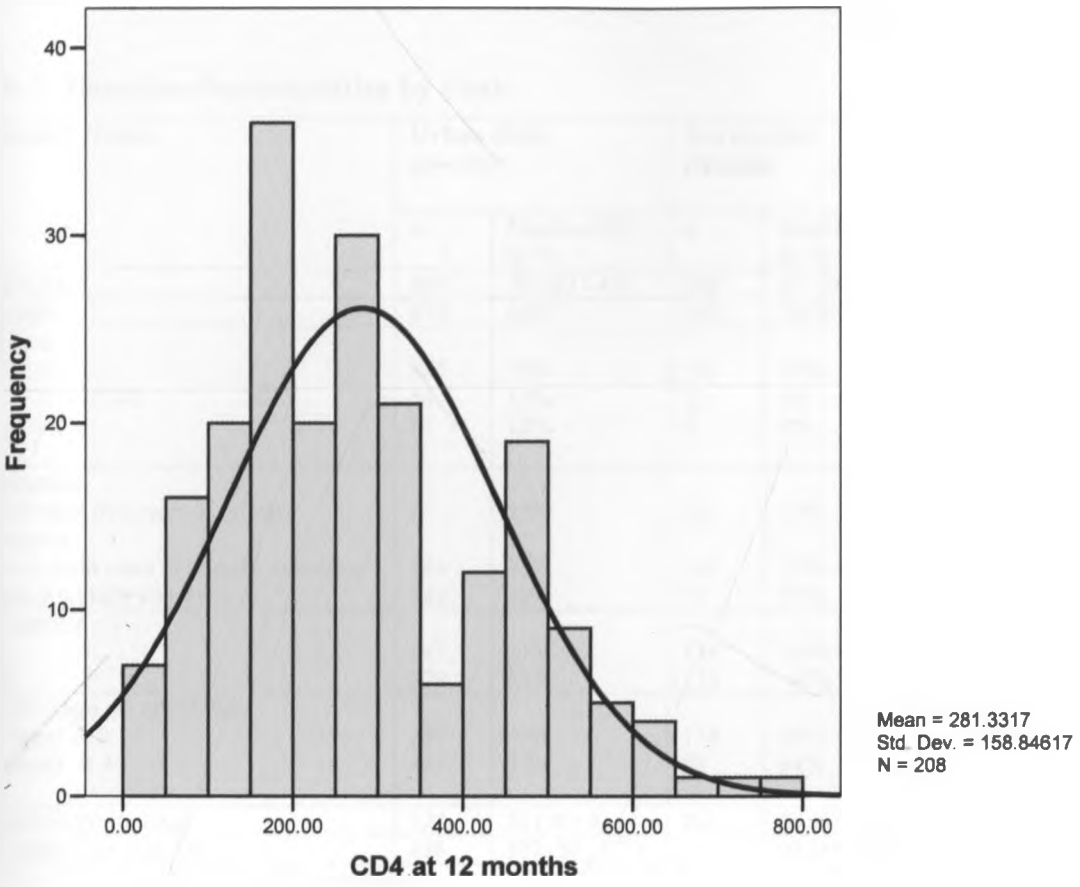


Chart 5: A histogram for cd4 after twelve months in the clinic



The patient's Cd4 and weight are normally distributed as evidenced by the above char

Some factors are said to be associated with mortality. This is according to similar studies done in the past as discussed in the literature review. These factors include age, gender, income, level of education, disclosure status WHO stage on enrollment for treatment, CD4 and weight on enrollment for treatment. These factors are independent of the type of clinic one is enrolled in.

Preliminary analysis was done on these factors. I compared the clinical and socio-economic characteristics for the HIV infected patients from the two settings.

Table 1 has the outcome of the analysis.

Table 1: Baseline characteristics by clinic

Characteristics	Urban clinic (N=456)		Rural clinic (N=268)		P value
	n	Median (IQ) or %	n	Median (IQ) or %	
Age(yrs)	456	35 (30 - 42)	268	37 (29 - 44)	0.001
Female	270	59%	160	60%	0.9
Income					
< 5,000	325	76%	215	87%	
5,001 – 10,000	54	13%	22	9%	0.0026
> 10,000	51	12%	9	4%	
Education					
Lower than five years of primary education :	45	10%	33	13%	0.003
Five to eight years of primary education:	164	38%	141	57%	
Beyond primary education :	224	52%	75	30%	
Disclosure					
No:	167	37%	134	52%	0.001
Yes:	279	63%	123	48%	
WHO stage on enrollment:					
Stage1 & 2	287	63%	173	66%	0.375
Stage3 & 4	167	37%	88	34%	
Baseline Weight(kg)	428	55 (48 - 63)	268	53 (47 - 59)	0.002
Baseline Cd4 (cell/ul)	436	107 (50 - 177)	219	92 (45 - 163)	0.053

IQ: 25th – 75th percentile

There is significance statistical difference between the two clinic set ups in the patient's income levels, education levels, disclosure status and the baseline weight ($p<0.005$). The group in the urban setup has better education which leads to better income. They are therefore in a position to afford better meals, can buy medications for the opportunistic infection, are updated on quality HIV care thus having better weight as compared to those in the rural setup.

The (p=0.375) and (p=0.053) for the WHO stage and baseline CD4 respectively indicate that there is no statistical difference between these indicators across the two setups. This is due to the fact that people tend to seek treatment when they feel sick regardless of the setup in which they are in. The two indicators help in determining how far the disease has progressed in one's body.

There is a lot of stigma associated with HIV/AIDS in the society today. This has led to the HIV patients not disclosing their status for fear of discrimination. The p value ($p<0.005$) indicates that the level of stigma is different in the two setups. A high percentage of those in the rural area do not disclose since the level of stigmatization is high in the rural setup. Disclosure is a sign of one having accepted his / her status and living positively. This in turn leads to the person on being in a position to seek services that are beneficial to the management of the disease

Table 2 (*Follow up Characteristics by clinic*)

Characteristics	Urban clinic (N=456)		Rural clinic (N=268)		P value
	n	Median (IQ) or %	n	Median (IQ) or %	
Adherence (*self reported*) Ever missed more than half the drugs Never missed more than half the drugs	21 404	5% 95%	39 217	15% 85%	0.0037
Month 6 CD4	257	244 (158 - 346)	101	178(112 -271)	0.002
Month 12 CD4	153	283 (175 – 447)	55	177 (132 – 276)	0.001
Month 6 Weight	345	58 (51 – 64)	175	56 (51 - 62)	0.047
Month 12 Weight	204	61 (54 - 68)	100	57 (51 – 63)	0.013

IQ: 25th – 75th percentile

There are some characteristics that influence how soon a patient may die of HIV/AIDS and may be influenced by the type of clinic the patient is getting care from. Table 2

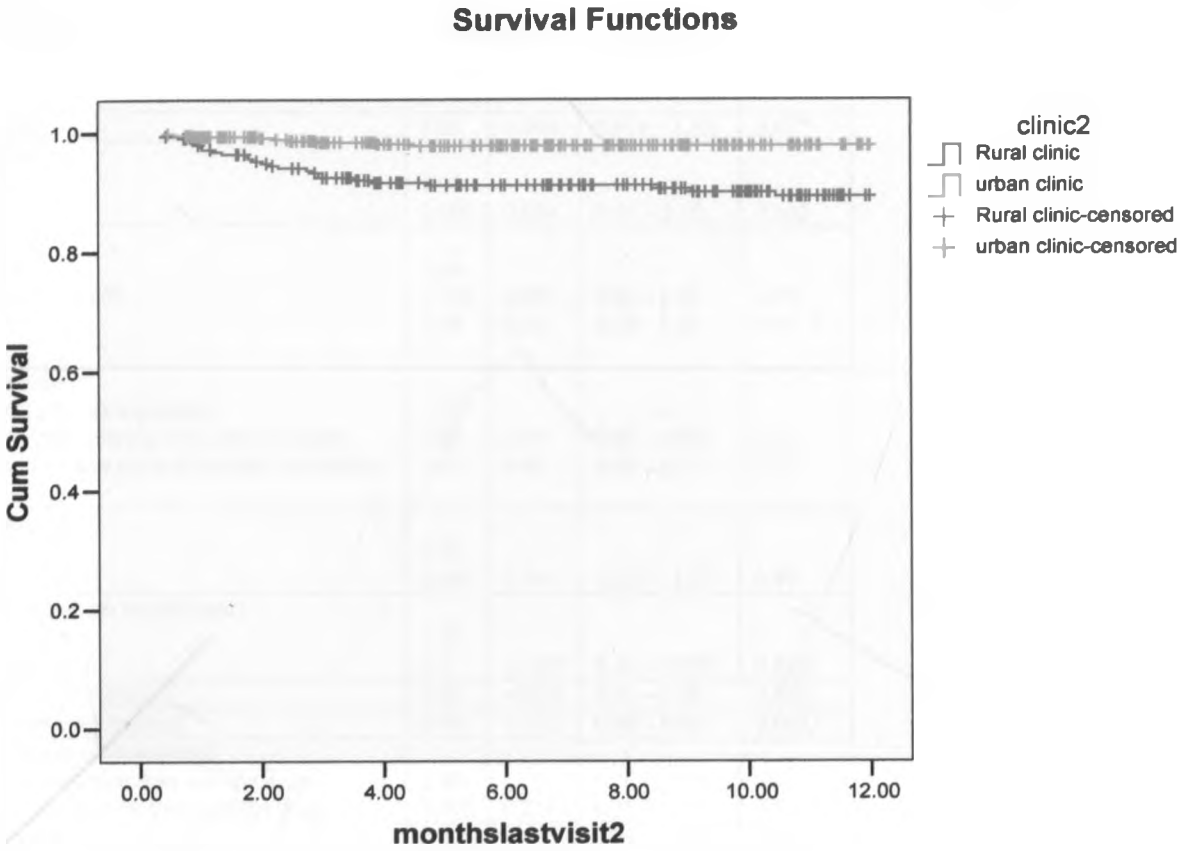
shows the summary of such factors. There is a significant difference between these factors across the two clinic settings under study ($p<0.005$) for all the factors.

The quality of service given in a clinic, which is determined by the available resources, play a big role in adherence to drugs by the patients in the clinic.

Month6 and month12 CD4 and month 6 and month 12 weight are the indicators that measure the improvement or deterioration of the patient's health on being initiated on ARVS. There is a significant difference in these two factors across the two settings. The patients in the urban setting have better CD4 at month 6&12, median (244) and median (283) respectively as compared to those in the rural setting with a median (178) and median (177) for month 6&12 respectively.

A gain in CD4 leads to an eventual gain in weight. There is an increase in weight from baseline weight to month6 weight and month 12 weight for patients in both settings.

Graph 1: Kaplan Meier graph



Graph1 is a plot of the survival functions in the two clinic setups.

The (p 0.001) for the survival functions indicate a significant difference between the mortality rates in the two settings. Mortality rate in the rural setting is higher than in an urban setting.

Table3: Univariate analysis

Characteristics (*Exposures)*	HR	β	95% CI	P value
Clinic				
Urban Clinic	1.00		-	
Rural Clinic	5.27	1.66	2.38 -11.7	0.001
Age(yrs)	1.01	0.009	0.974 – 1.05	0.617
Gender				
Female	1.00		-	
Male	1.093	0.089	0.55 – 2.18	0.802
Income				
> 10,000	1.00		-	
5,001 – 10,000	1.102	0.097	0.61 – 1.99	0.75
< 5,000	1.24	0.22	0.22 – 1.82	0.45
Education				
Beyond primary education:	1.00		-	
Five to eight years of primary education:	1.28	0.25	0.87 – 1.88	0.21
Lower than five years of primary education :	1.51	0.41	0.84 – 2.74	0.17
Disclosure				
Yes:	1.00		-	
No:	0.986	0.014	0.563 – 1.97	0.97
WHO stage on enrollment:				
Stage3 & 4	1.00		-	
Stage1 & 2	0.5	-0.698	2.49 – 0.997	0.049
Baseline Weight(kg)	0.94	-0.059	0.91 – 0.98	0.002
Baseline Cd4 (cell/ul)	0.99	-0.012	0.98 – 0.99	0.001
Adherence (*self reported)*				
Ever missed more than half the drugs	1.00		–	
Never missed more than half the drugs	1.290	0.255	0.31 – 5.46	0.73
Month 6 CD4	1.00	0	0.99 – 1.01	0.979
Month 12 CD4	0.975	-0.026	0.94 – 1.02	0.225
Month 6 Weight	0.944	-0.058	0.869 – 1.025	0.168
Month 12 Weight	0.992	-0.008	0.87 – 1.132	0.908

Cox PH model was fitted for those factors that were thought to have an influence on mortality. Statistical significance of each variable will be assessed to decide on the association between the independent variables and AIDS mortality rate.

Table 3 shows the results of the analysis.

Type of clinic: The risk of mortality for those in the rural clinic is 5 times higher than those in the urban clinic HR(5.27) and is statistically significant p(<0.001)

Age: For every additional year in age, the risk of mortality decreases by 1% HR(1.01) though not statistically significant .

Gender: The risk of mortality among the male is 9% higher than that of the female HR(0.802) though not statistically significant.

Income: The risk of mortality among those earning Ksh5,000 -10,000 is 10% higher than that of those earning >Ksh10,000 HR(1.102) while for those earning <KsH5000 is 24% higher HR(0.45) though not statistically significant for all.

Education: The risk of mortality among those who have learnt 5 – 8yrs of primary education is 28% higher than those who have learnt beyond primary education while it is 51% higher in those with less than 5yrs of primary education relative to those who have learnt beyond primary education. HR(1.28) and HR(1.51) respectively though not statistically significant.

WHO stage: The risk of mortality among those in stage 1&2 is 50% lower than that of those in stage 2&4 though not statistically significant.

Weight: The risk of mortality decreases with every unit increase in weight. This is statistically significant in baseline weight but not in month 6&12 weight.

Weight: The risk of mortality decreases with every unit increase in weight. This is statistically significant in baseline weight but not in month 6&12 weight.

The type of clinic, the baseline cd4 and the baseline weight were associated with mortality.

Table 4: Multivariate Analysis

	HR	95% CI	P value
Model 1			
Baseline Cd4 (cell/ul)	0.99	0.98 – 0.995	<0.001
Clinic Rural clinic Urban clinic	1 0.2	- 0.089– 0.45	<0.001
Model 2			
Baseline weight (kg)	0.95	0.103 – 0.51	<0.001
Clinic Rural clinic Urban clinic	1 0.952	- 0.92 – 0.99	<0.001

Cox PH model was used to determine the factors associated with mortality. Table 4 shows the results of the analysis. A model with either baselineCD4 or baseline weight and the clinic were the best models

4.2 Conclusion and recommendations

There is a difference in mortality rates for HIV patients in an urban setting and those from a rural setting.

Though there is no significance difference in how sick the patient is at the point of seeking treatment, there is a difference in response to treatment for the two groups. The group in the urban area responds better than the rural group.

This project did not bring out the relationship between HIV/AIDS prevalence and AIDS mortality rate. Earlier studies have proved that an increase in HIV/AIDS prevalence in the urban areas does not increase the mortality rate but increase prevalence leads to an increase in mortality rate in the rural areas. This shows that there are ways in which death is delayed in the urban areas which can be applied in the rural areas too.

There is need to evenly distribute health facilities across the country. Currently most of the health facilities are located in the urban areas.

Definitions

CD4 – It helps to tell the doctor how strong your immune system is, how far HIV disease has advanced and helps to predict the risk of complications. The count goes down as the disease progresses.

ARVs – These are medications taken for treatment of infection by HIV. It is usually a combination of two or more drugs.

Viral load – It measures the level of HIV in the blood to determine the staging and outlook of the disease.

ARV naïve –The patient has never taken ARVS before

Eligible for ARV – The patient is eligible to start taking ARVS according to clinical assessment

Who stage – World Health Organization staging

References

1) Berry, McKinney and Mcclain,1996,carwen, Sabo Berry 1993, Graham, Forrester, Wysong, Rosenthal & janes 1995, voismer 1992.

2) Buehler, chu and the AIDS mortality group, 1995, David & Stapleton, 1991, Rumley & Eisinhart, 1993

3) CDC report 1993

4) Egger M, F dabis and M schechter. Mortality oh HIV-1-Infected patients in the first year of antiretroviral therapy: Comparison between low-income and High income countries, *Lacent 2006; 367:817-24.*

5) Family life counseling association of Kenya, 1996.

6) Timothy Lahey, Michelle Lin, Bryan Marsh, Jim Curtin, Kim Wood, Betsy Eccles, C Fordham Von Reyn. Aids Research and Human Retroviruses. May 1,2007, 23(5):963-698. doi:10.1089/aid.0206

7) Kallings LO (2008). "The first postmodern pandemic: 25 years of HIV/AIDS". *J Intern Med* **263** (3): 218–43.

8) UNAIDS, WHO (December 2007). "2007 AIDS epidemic update" (PDF). http://data.unaids.org/pub/EPISlides/2007/2007_epiupdate_en.pdf.

9) Kenya ministry of health, 2001 sentinel surveillance system, Aids in Kenya,.

10) Samuel Adari phd. Poverty exacerbates HIV/AIDS mortality by Johnson.

11) United Nations millennium development goals 2006.