Understanding Disease Progression in the Kagera Region of Tanzania: A Framework for Efficient Health Care Delivery

A Thesis
submitted to the Graduate School of Arts & Sciences
at Georgetown University
in partial fulfillment of the requirements for the
degree of
Master of Public Policy
in the Georgetown Public Policy Institute

By

Sarah E. Berger, B.A.

Washington, DC
April 14, 2008

UNDERSTANDING DISEASE PROGRESSION IN THE KAGERA REGION OF TANZANIA: A FRAMEWORK FOR EFFICIENT HEALTH CARE DELIVERY

Sarah E. Berger, B.A.

Thesis Advisor: Marcela Tarazona, Ph.D.

ABSTRACT

Today, chronic illnesses remain a great health burden for households and individuals in sub-Saharan Africa. This has been especially true in the Kagera Region in Tanzania, where HIV/AIDS rates have historically been greater than the national averages and health services have been poorly administered. In Kagera and elsewhere, policymakers must understand the progression of specific diseases in order to develop appropriate guidelines for treatment and care, and prioritize and allocate resources for health services. This study attempts to provide such a framework. Using longitudinal data from the Kagera Health and Development Survey 1991-1994, and 2004, the study examines health patterns among adults (aged 15-95) in Tanzania, and explores how well chronic health symptoms explain variations in survival time. This analysis shows that there are a number of populations that are at higher risk of experiencing death in the Kagera region: individuals with persistent weight-loss; females; individuals aged 35-44 years of age; previously married men and women; and individuals with higher levels of education. Future research is needed to explain why it is that some parts of the population survive less relative to individuals without those characteristics.

The research and writing of this thesis is dedicated to all those who supported my work: Marcela Tarazona, thesis advisor; Jeff Mayer, mentor and writing tutor; Kathleen Beegle, Senior Economist at the World Bank, mentor and reader; Shwetlana Sabarwal, Economist at the World Bank; Lucia Fort, Senior Gender Specialist at the World Bank; A. Waafas Ofosu-Amaah, Senior Gender Specialist at the World Bank; and the Berger gang.

TABLE OF CONTENTS

Tables

Figures

I. Introduction

Disease burden and chronic illnesses remain alarmingly high in sub-Saharan Africa. Many countries in the region have experienced significant levels of HIV/AIDS and other chronic illnesses that are potentially contributing to demographic, economic, and social changes affecting future growth and development. At the end of 2001, in sub-Saharan Africa, nearly 7.7 million individuals, predominantly infants (0-4) and adults (aged 15-49) were burdened with chronic illnesses (Mathers, Lopez and Murray, 2006). In Tanzania, in particular, chronic illnesses and other communicable diseases have had an overwhelming affect on the adult population. In 2005, the World Health Organization's World Health Statistics Report for Tanzania indicated that more than 75 percent of the disease burden in that country is a result of communicable diseases like HIV/AIDS, malaria, tuberculosis, and diarrhea (World Health Organization, 2005).

Tanzania's experience has been especially severe. In Tanzania today, the adult mortality rate for adults (15-60) remains high. The probability of death for males, aged 15 to 60 years is reported at 541 deaths per 1,000 individuals and females, 505 deaths per 1,000 individuals. Life expectancy for males is 47 years and for females in 49 years (World Health Organization, 2006). HIV/AIDS is partly to blame. In 2002, 29% of all deaths in Tanzania were estimated to be HIV/AIDS related (World Health Organization, 2006). In the region of focus of this paper, rates of HIV/AIDS in the early 1990s were estimated at 28.6% among women and 16.9% among men (Mshana, et al., 1998). Persons with the greatest incidence were located in the urban areas of the Kagera Region and among females, aged 15-24 years-old (Killewo, et al.). More recent studies have shown prevalence rates to be substantially lower.

In Tanzania as elsewhere in the sub-Saharan region, insufficient health care has compounded the affects of chronic diseases. In 2005, the WHO reported that while Tanzania has had “adequate and equitable” coverage of health care facilities, quality of care has generally been low and staff is unequally distributed. Supplies of drugs and medical supplies have been inconsistent and scarce largely due to the country’s reliance on foreign assistance for health services funding. Finally, limited capacities in producing reliable data and analysis have created a strain on the country’s health information system (World Health Organization, 2005). Without access to pertinent medicines or necessary health data, prevention of curable communicable diseases has been difficult.

Up to now, implementation of targeted, preventative adult health care policies in sub-Saharan Africa has relied heavily on data from verbal autopsies, hospital records, and some epidemiological studies (Mathers, Lopez, and Murray, 2006). In Tanzania, similar methods have been used to determine the impact of specific diseases. While data are slowly becoming more available, there remains a significant gap between the data and policy design and implementation. This study uses data on chronic illnesses to help move the policy discussion forward on efficient health care delivery.

Using longitudinal household data from the Kagera Region of Tanzania collected in 1991-1994 and 2004, this paper analyzes how well chronic health symptoms recorded in 1991 predict the duration of time until death for prime-age adults. The study focuses on the four chronic health symptoms that are reported in the Kagera Development and Health Survey (KHDS): chronic diarrhea, persistent fever, rash for more than a month, and weight-loss. These chronic health symptoms are strongly associated with AIDS status of adults (Eramova, Matic, Munz, 2006). This model does not try to attribute deaths specifically to AIDS but recognizes that it is a major cause of death in the region.

The next section reviews the literature on chronic illnesses in developing countries and methods used to for measuring the duration of time until death. The third section describes the analytical framework that justifies the use of key explanatory variables in the study's model analytic. The fourth section presents the dataset, the methodology to be used in the multivariate analysis, and the descriptive statistics. The fifth section presents the results of the multivariate analysis. The final section is a discussion of the results.

II. Literature Review

Chronic and communicable diseases

Chronic and communicable diseases are frequent causes of mortality and morbidity among adults in sub-Saharan Africa. In 2001, the World Health Organization collaborated with the World Bank and the Harvard School of Public Health on the

Disease Control Priorities Project to determine what the most burdensome diseases were and how policies, funds, and energies could be more effectively used to combat them.

Mathers et al. (2006), using data from sub-Saharan Africa confirm that communicable diseases including HIV/AIDS, tuberculosis, and diarrhea remain some of the most frequently cited illnesses among adult men and women in the region. Lopez, Begg and Bos (2006) support the findings of Mathers et al., and add that communicable diseases in sub-Saharan Africa are not evenly distributed across gender, age, and location. In general, they say communicable diseases disproportionately affect women and youth (aged 15-24 years). While the authors do not explain why this is the case, I plan to explore this finding more completely using panel data from Tanzania.

Definitions of major and minor signs of chronic illnesses coincided in this study are drawn from the World Health Organization's classification of AIDS-related symptoms. Symptoms noted include persistent fever, chronic diarrhea, extreme weight-loss, and a prolonged rash (WHO, 2007). Eramova, Matic, and Munz (2006) elaborate on these general chronic symptoms. They define persistent fever in adults is a body temperature of 38 degrees Celsius (100.4 degrees Fahrenheit) for more than two weeks. They also note that a persistent fever may be a symptom of other serious diseases including pneumonia, tuberculosis, or lymphoma. They define severe weight-loss as "the involuntary loss of more than 10 percent of one's body weight." Finally, they explain chronic diarrhea to be either episodes of acute diarrhea or three or more consecutive loose stools over 28 days.

Previous research and methodologies

Previous research and methodologies on common chronic and communicable diseases among adults in sub-Saharan Africa rely principally on the use of verbal autopsies. Verbal autopsies are data that describe a deceased individual's symptoms prior to death and are collected by lay personnel from a relative or friend of the deceased (Doctor and Weinreb, 2003a). This technique has been helpful in understanding the patterns and symptoms of different diseases, and how to allocate resources for patient care and the planning of intervention programs.

In 1997, Boerma et al. used longitudinal data on the causes of adult mortality in rural Tanzania to understand the most common diseases in the country. They concluded that the primary causes of death among adults aged 15-59 years were HIV/AIDS, chronic diarrhea, malaria, and cardiovascular disease. In Uganda, a similarly designed study (Kamali, et al., 1996) examined the magnitude of the HIV-associated mortality burden on a population. The authors surveyed family members about the signs and symptoms during the dead individual's last illness and probable cause of death, and compared the results with HIV-1 serology data obtained from annual repeat surveys of the same population prior to death. The authors confirmed that HIV/AIDS was the main cause of death, while there were only a small percentage of deaths that were not HIV-related.

Finally, Doctor and Weinreb (2003) conducted a study in Malawi to estimate AIDS-related adult mortality. Examining verbal autopsy data, they used the World Health Organization's classification of AIDS-related symptoms to determine whether individuals were HIV-positive. They confirmed that the WHO's classification tools were highly predictive of who had HIV/AIDS. However, while it is highly possible that this study has some external validity, I will not be applying these conclusions to my study.

Overall, verbal autopsies provide information about the frequency of a chronic illness or communicable disease; yet, the method is very limited. It reveals very little, for example, about the duration of the disease, the variation within a chronic illness, or the predictability of time until death. These details are important because they inform efforts to design appropriate health care plans that are illness-specific and address the most vulnerable population.

Assessments of the probability of death are easier to use if data are available prior to the individual's passing. Probability of death analysis gives weight to certain characteristics that may be more likely to predict death than others. Using data from a longitudinal study conducted in Zambia, Chapoto and Jayne (2005) employ a probit model to explain the relationship between the individual demographics of sick and non-sick persons (e.g. income, education, age, gender, predicted AIDS status, and access to health services) and their likelihood of dying. A unique element of this model is the creation of an HIV/AIDS variable. Chapoto and Jayne design an HIV/AIDS variable based on the WHO classification: two of the chronic symptoms and one minor symptom. This is extremely relevant for this present study because the data from Tanzania do not have a variable for HIV/AIDS. This study will employ Chapoto and Jayne's method of creating a proxy variable for HIV/AIDS to try and see if having at least two chronic symptoms decreases an individual's survival time.

Overall, using a number of specifications, the authors conclude that women are disproportionately afflicted by diseases. They also show that income and wealth are unlikely predictors of death, whereas the varying health symptoms have significant explanatory power. Finally, they find that prime-age men and women experiencing a prior death in their household are 23.0 and 18.1 times more likely, respectively, to die of disease-related causes than men and women in households with no prime-age adult deaths.

Doctor and Weinreb (2003b) also compare data from Kenya and Malawi to estimate the annual probability of dying among married men (age 20-59) in high and moderate HIV-prevalent rural areas. The authors employ standard life table calculations, a method that is frequently used to predict future impacts on mortality based on the changes in death rates (Miller and Hurley, 2003) and conclude that there are absolute differences between mortality in Kenya and Malawi. They also note that there is definite variation across age groups. Both the Chapoto and Jayne (2005) and Doctor and Weinreb (2003b) studies show that adult mortality has not only increased significantly, but that there are specific groups that are more vulnerable to dying. This information is helpful for efficient and well-targeted policy design. Nevertheless, the measurement of the dependent variable is difficult and explains very little about the duration of time until an individual dies.

Survival Analysis

A more common methodology employed in the field of epidemiology is survival analysis. Survival analysis, sometimes referred to as duration analysis, analyzes the time to failure, where ‘failure’ signifies a single discrete event (Gutierrez, 2007). In the case of this present study, the event is death. Survival analysis commonly uses panel and cohort surveys with repeated interviews on a sample of interest (Jenkins 2005). Unlike logit and probit models, the dependent variable is the “waiting time until the occurrence of a well-defined event” (Rodríguez, 2007). A unique element of a survival analysis model is that it is able to retain observations for which the event has not occurred and then take the effects of these observations into account mathematically (Reed et al., 1994). In the context of health policy, survival analysis is a helpful tool for determining the duration of time an individual is likely to survive with certain health symptoms and other demographic characteristics.

Epidemiological studies have applied survival analysis to examine the time of survival until death. Reed et al. (1994) use survival analysis to examine gay men with AIDS and how well realistic acceptance or knowledge of the disease predicts the duration of survival. The authors assess the difference between those who had a “low” realistic acceptance and those who had a “high” realistic acceptance and find that those who realistically accepted that they were ill experienced a significant decrease in survival time. Similarly, Zangerle et al. (1995) use survival analysis to determine the effect of clinical and demographic factors on survival among 901 AIDS cases in Austria. They conclude that there are considerable differences in survival time between different HIV-risk groups (intravenous drug users, heterosexual contact, and homosexual contact).

In a more recent study in Brazil, Braga et al. (2007) use survival analysis to look at the effects of gender differences on survival in HIV/AIDS cohorts. Conditional on the fact that men and women have equal and free access to ARTs and other drugs that alleviate the mortal effects of HIV/AIDS, they find survival rates to be dramatically different between males and females (holding all other variables constant). In one specification, being female was shown to be a good predictor of shorter survival. Zhou and Kumarasamy (2005) also employ survival analysis in Asia and the Pacific region to predict short-term disease progression among HIV-infected patients, half of whom were receiving antiretroviral treatment and the other half who remained as a control group. They report that, compared to patients receiving antiretroviral treatment, patients not on treatment had a higher rate of disease progression. While all of these studies differ in their employment of survival analysis, they draw notable conclusions about the length of time a person with certain characteristics survives with a particular disease. This, in turn, lends to specific policy implications.

Most of these studies look at the detrimental impact of HIV/AIDS. While the Kagera region reportedly had one of the highest HIV/AIDS rates in Tanzania at the time of the baseline survey, it is unreasonable to assume that all deaths are due to HIV/AIDS. As a result, the present study will differ from previous studies on HIV/AIDS and, instead, explore how well AIDS symptoms (chronic fever, persistent rash, chronic diarrhea, and significant weight-loss) or other possible health demographics predict duration of survival.

Policy Relevance

Understanding how well chronic health symptoms actually predict the time a person survives has a number of policy implications. First, it has implications for how well a nation's health policy targets the most vulnerable individuals. Well planned gender-, age-, and illness-specific health policies are the most effective and efficient way to serve the most vulnerable populations. Survey data may show that being female and having a chronic illness, holding all other variables constant, is likely to increase survival time compared with males. If this is the case, then policymakers would have to explore whatever it is about males that make them have lower survival times; and more importantly, to decide what policies need to be designed and implemented to address this health issue.

Using survival analysis as a tool also has its own policy implications. Good health care policy takes into account the duration of a disease. Survival analysis provides estimates of the probability of disease progression, which, in turn, helps policymakers to develop appropriate guidelines for treatment and care, and to prioritize and allocate resources for health services, especially in developing countries (Zhou and Kumarasamy, 2005). This is particularly relevant to Tanzania, where medical supplies and resources are scarce. Understanding survival patterns for individuals with chronic health symptoms will contribute to well-planned effective health care policies.

III. Conceptual Framework and Hypotheses

Using data from the KHDS, this paper analyzes the determinants of life expectancy as a function of exogenous individual characteristics (age, gender, education, migration, and marital status) and specific health factors (chronic illnesses, duration of illness, visit to a doctor). The units of analysis are individual people. Because of the statistical methodology employed, I have narrowed my sample to household members who were interviewed in the first round of the KHDS and reported a chronic illness between September 1991 and May 1992. Efforts were made in 2004 to re-contact these individuals. Thus, the data capture information about individuals in the sample at two discrete time periods, and whether they are alive or dead at the time of the last interview attempt. Overall, though, I examine individuals who reported at least one of the four chronic illnesses in 1991 and estimate their relative risk of survival between 1991 and 2004.

This study uses survival analysis, or 'time to event analysis,' where the focus is on analyzing the time leading up to an event. For this analysis, the event is defined as death. There are two specific elements of survival analysis data, censoring and non-normality, which do not allow for the use of traditional statistical models such as multiple linear regressions (UCLA Academic Technology Services, 2008). First, the non-normality aspect of the data implies that the sample error is not normally distributed with zero mean and a constant variance. This violates one of the classic linear model assumptions (Wooldridge, 2006). The other element unique to survival analysis is censoring. The UCLA Academic Technology Services defines a censored observation as an observation (or individual) with incomplete information. This study uses data that are right censored, meaning that the information is incomplete because the individual did not experience a death during the time-span of the study (UCLA Academic Technology Services, 2008). The purpose of using survival analysis for this study is to follow subjects over the 13 years and to observe at which point in time they experience death, if at all.

This study employs the proportional hazards model to incorporate the specific elements of survival analysis. The proportional hazards model describes how hazard (or risk of death) changes over time. In particular, it calculates the hazard rate which is the propensity of an individual to die at a point in time, given that the individual has survived up until that point of time. The hazard rate is the fundamental dependent variable of survival analysis (UCLA, 2008). It reflects three main elements: a begin time; an end time; and a failure/non-failure indicator or occurrence (Gutierrez, 2007). In this study, the begin time is the age of the individuals in the first round of interviews in 1991; the end time is the age of the individuals at time of death or in 2004; and the failure/non-failure indicator is death, where individuals who did not die did not record an occurrence.

In order to make statistical inference, the analysis uses the hazard rate in the Cox proportional hazard regression model. The Cox regression model is a hazard function that calculates the hazard ratio, which is the rate of a group of individuals that experience death at time t over the hazard rate of the baseline group all the individuals that have survived up to and including that time period (Yaffee, 2001). The assumption associated with the proportional hazards model is that the ratio of hazards is constant over time across groups. For example, the calculated hazard ratio of the risk of death for males compared to females in the Kagera region is assumed to be constant between 1991 and 2004. Overall, a higher hazard ratio is associated with shorter survival, and a lower hazard ratio is associated with longer survival.

Using the Cox proportional hazards regression model, this study explores three different hypotheses.

Hypothesis 1: An individual reporting a chronic illness is predicted to survive longer if he/she has visited a doctor and/or experienced chronic symptoms for less than a year.

In 1991, when baseline survey data were collected, the Kagera region of Tanzania had limited access to resources such as medical supplies and doctors. Between 1991 and 2004, the situation improved only minimally (Belgian Development Cooperation, 2004). In effect, the region during that time period was insulated from other exogenous influences. Hypothesis 1 examines how the variation of independent health variables affects the survival time of individuals experiencing a chronic illness. The risk of survival for individuals with chronic illnesses who have visited a doctor is expected to be lower because a doctor is assumed to provide the medicines and/or care necessary to ameliorate the effects of an illness. Individuals who do not visit the doctor, however, are expected to have a shorter survival time. Not being able to visit the doctor may also be a result of not having the financial means to seek treatment or access to care. This is controlled for in this study.

Duration of the illness is also expected to have adverse effects on individuals reporting a chronic illness. Individuals who have experienced a chronic illness or one of the four chronic health symptoms measured in the survey (fever, chronic diarrhea, extreme weight-loss, skin rash lasting for more than a two years) are expected to have a higher risk of death relative to those who only recently report an illness or symptoms.

Overall, these unique health conditions are predicted to have a significant effect on how long an individual will survive.

A second hypothesis looks at how well variations in demographics (sex, age, education, migration, marital status) affect the survival time of an individual with a chronic illness.

Hypothesis 2: Individuals with reported chronic illnesses are likely to have shorter survival times if they are less educated, previously married or never married, were born outside the Kagera region, are between the ages of 25 and 35, and are male.

In a number of studies, education has been shown to increase an individual's awareness of health issues and possibly change risky behavior (Sinha and Kiso, 2007). While these factors have little apparent effect in the short run, they have been shown to have significant effects in the long run. Individuals with chronic illnesses and higher levels of education may have better knowledge of how to access medical care or treatment.

This hypothesis also predicts that males will survive for less time than females. As previously mentioned, the World Health Organization reported that the average life expectancy for males in Tanzania was about 47 years and about 49 years for females (WHO, 2006). While this country demographic may not capture the average life expectancy of males and females in the Kagera region, it portrays women in general as living longer than men.

Two other demographics that are likely to cause variation in survival time are migration and marital status. In many households in the Kagera region, migration for work is common, especially among men. Migration has been shown to increase one's likelihood of contracting HIV/AIDS, among other diseases (UNAIDS, 2007). As a result, individuals with at least one chronic symptom who have migrated are predicted to have shorter survival times. Finally, marital status may also affect an individual's survival time. If an individual is single, widowed or divorced, he/she may find it difficult to gain the support or care necessary for longer survival; whereas, those with chronic symptoms who are married may have stronger social networks and therefore more individuals to care for them.

A final hypothesis examines the effects of specific HIV/AIDS symptoms:

Hypothesis 3: Individuals with two or more chronic health symptoms are assumed to be HIV-positive and therefore experience a shorter survival time than those individuals only reporting one chronic illness.

In a study of the characteristics of individuals afflicted by AIDS-related mortality, Chapoto and Jayne (2005) use data from Zambia to determine the probability of someone dying based on the relationship between specific characteristics. They classify an "AIDS-related" death as an individual who exhibited at least two major signs of HIV/AIDS (major signs are defined by the World Health Organization as rapid weight-loss, chronic diarrhea, rash, or persistent fever). The present study employs the same method for classifying individuals who are *potentially* HIV-positive.

On average, HIV-positive individuals who do not receive antiretroviral treatment are predicted to live 1 to 2 years from the onset of the virus (WHO, 2006). As previously mentioned, between 1991 and 2004, ARV treatment was not available in the Kagera region to help slow the progression of HIV/AIDS (Belgian Development Cooperation, 2004). Under Hypothesis 3, individuals who report two or more chronic symptoms are assumed to be HIV-positive, and are expected to have shorter survival times compared to those individuals reporting only one chronic illness.

IV. Data and Methods

Data

The KHDS is a longitudinal living standards survey designed, implemented and published by the World Bank, and intended to examine the socio-economic impact of the death of prime-age adults on other household members (Ainsworth, Over, et al., 2004). The survey instrument includes sections that capture different socio-economic factors including education, health, consumption, and crop production.

The present study uses data gathered in the health section of the survey, which surveys individuals on acute and chronic illnesses. It is important to note that this survey did *not* attempt to measure knowledge, attitudes, behaviors or practices with regard to

HIV/AIDS. Blood samples were not drawn and HIV sero-prevalence was not tested. As a result, information is limited to what was verbally reported.

The survey over-sampled sick households to ensure that the data would capture the impact of death or dying adults. Between March and June 1991, researchers identified 51 communities as primary sampling units (PSU's). Within those PSU's, 29,602 households were surveyed on whether or not any adults in the household were ill at the moment and unable to work, or if an adult aged 15-50 had died within the past twelve months. In the end, researchers labeled 27,433 households as "well" and 2,169 as "sick." From this information, households were stratified according to the degree of adult illness and mortality. Sixteen households were selected from each PSU, fourteen which were randomly selected from the "sick" (*type A)* households and two which were randomly selected from the "well" (*type B)* households. In one PSU, the number of "sick" households was less than 14, so all of them were included and the numbers were balanced with well households (Beegle, De Weerdt, Dercon, 2006b).

To reduce the effect of attrition on survey results, the KHDS was conducted in four distinct time intervals, or *passages*, 6-7 months apart over the period 1991-1994. The first passage of fieldwork occurred between September 1991 and May 1992; the second passage between April 1992 and November 1992, etc. During each passage, questionnaires were administered once in all the households.[1]

Within each passage, there were also a range of questionnaires referred to as *waves*. The *wave* assigned to any particular household was dependent on which questionnaire the household had previously answered. Each *wave* of questionnaires asked the same questions, but with some slight variations. All households interviewed for the first time (regardless of what passage) were given the *wave 1* questionnaire; those interviewed for the second time received a *wave 2* questionnaire, and so forth. When a household dropped out mid-survey, it was replaced with another household, which started with the *wave 1* questionnaire, regardless of the passage number. Thus, the wave number does not correspond with the passage number (Ainsworth, Over, et al., 2004).

The KHDS began with 816 households in the first passage of interviews. At the conclusion of the fourth wave, 759 households had completed all four waves, 69 had completed three waves, 45 had completed 2 waves and 39 households had completed only one wave. The net number of households interviewed was 912 households (totaling to 6,204 individuals) (Ainsworth, Over, et al., 2004 and Beegle, De Weerdt, Dercon, 2006b). Individuals over the age of 7 years responded to the survey themselves, while heads of households responded for children younger than 7 years. Overall, data were collected on over 6,000 individuals in the Kagera region between 1991 and 1994.

[1] This survey also collected community level data on local markets, school and health facilities, and local

In 2004, World Bank researchers re-contacted and re-interviewed all the households who participated in *any* round of the KHDS in 1991-1994 and whose members were alive during the last interview in 1994. The KHDS 2004 also collected information about those household members living during the last interview, who had passed away by 2004. KHDS 2004 achieved a 93 percent rate of re-contact[2] with the baseline households. In 832 of the 912 households, at least one household member was re-interviewed.

The present study will focus on household members interviewed in the first round of the KHDS whose members were re-contacted in 2004. The analysis uses this cohort to determine how well reported chronic health symptoms predict the survival time.

Analysis Plan

The survival function is often noted as S (t) = Pr {T > t}, where *T* is any continuous random variable including a variable for whether the event failed; t is the time at which the event occurred; Pr is the probability that T is greater than t; and S (t) is the estimated probability of surviving at time t, given the other predictors, T (Yaffee, 2001). This equation gives the probability that a household member has not died by the end of duration t (Rodríguez, 2007). Because there is no direct way to estimate the survivor function, the hazard function is commonly used to determine the probability of

---

leadership structures; but, the present study does not use these data (Ainsworth, Over, et al., 2004).
[2]Re-contact is defined as having interviewed at least one person from the household.

individuals living at a certain point in time. The hazard function, therefore, can be thought of as the mortality rate over a period of time. Thus a lower hazard ratio predicts a longer survival time. In the present study, survival times are predicted to vary greatly.

Effects of Health Indicators

The present study employs a number of specifications in order to address the different hypotheses. In the first group of specifications, the key health indicators are tested for how well they predict variation of survival times among individuals who reported having one of the four chronic symptoms. These specifications test the first hypothesis.

Independent variables include: the four AIDS-related chronic symptoms as dummy variables (chronic diarrhea, persistent rash, chronic fever, persistent weight-loss); a dummy variable indicating whether or not the individual has recently visited a doctor; dichotomous variables for different levels of duration of the chronic illness in weeks, with the baseline set at 0-26 weeks (1/2 a year). Body Mass Index (BMI) (to control for overall general health) was also broken down into three dichotomous variables to capture those individuals who are underweight, at a healthy weight, or overweight. Gender is measured with a dummy variable, *male*, where male=1 and female=0. Dichotomous variables are created for various age groups with the baseline set to 15 years of age.

(1) **S(t)=** Pr { β_1Chronic Diarrhea, β_2Chronic Fever, β_3Persistent Rash, β_4Persistent Weight-loss > t }

(2) **S(t)=** Pr{β_1Chronic Diarrhea, β_2Chronic Fever, β_3Persistent Rash, β_4Persistent Weight-loss, β_5BMI, β_6Duration of Chronic Health Symptoms, β_7Visit to Doctor, β_8Male, β_9Age > t }

Effects of other independent variables

An additional specification is run to determine how the effects of other demographic variables are at predicting an individual's survival time. Other demographic independent variables include years of education, household consumption, marital status, sex of head of household and years of education of head of household. The inclusion of these variables aims at testing the second hypothesis that individuals with reported chronic illnesses are likely to have a lower risk of survival if they are less educated, previously married or never married, are born in the Kagera region, are between the ages of 25 and 35, and are male.

Education is captured in four different dichotomous variables, where different grade levels are grouped to create one variable. *Education* is measured by the number of school years completed. *Born in the Kagera Region* is a dummy variable; if an individual migrated the variable=1, otherwise it is 0. Dichotomous variables are also created for different levels of household consumption. The variable *Marital Status* is coded as three dichotomous variables: one variable for never married; one variable for previously married; and one variable for married. Variables are also included for sex of the head of household and years of education for the head of the household. Finally, two variables for specific household characteristics (e.g. if the household has a toilet or if the household has a non-earth floor) are used to control for the individual's living environment. In this specification, these additional independent variables are predicted to add significant variation to the probability of surviving.

S(t)= Pr { β_1Chronic Diarrhea, β_2Chronic Fever, β_3Persistent Rash, β_4Persistent Weight-loss, β_5BMI, β_6Duration of Chronic Health Symptoms, β_7Visit to Doctor, β_8Male, β_9Age, **β_{10}Education, β_{11}Marital Status, β_{12}Born in Kagera Region, β_{13}Toliet, β_{14}House with Floor, β_{15}Sex of Head of Household, β_{16}Years of Education (head of household), β_{17}Total Household Consumption**> t }

Predicted HIV/AIDS Related Deaths

A third group of specifications are run with the synthetic *HIV/AIDS* variables. The created *HIV/AIDS* variables employ Chapoto and Jayne's (2005) method of defining HIV/AIDS (at least two major AIDS symptoms). Using a combination of the four aforementioned chronic illnesses, the paper constructs four synthetic variables based on the four symptoms in the study. The specifications below apply the same method and independent variables from above, however, with the created *HIV/AIDS* variable. HIV1 is defined as chronic fever and weight-loss; HIV2 is chronic diarrhea and weight-loss; HIV

3 is rash, diarrhea, and weight-loss; and HIV 4 is chronic fever, weight-loss and rash. These variables are discussed in greater detail later in the paper.

The created *HIV/AIDS* variables are predicted to have varied effects on the survival time of individuals with reported chronic symptoms. Household members with chronic weight-loss and chronic diarrhea are predicted to have a lesser probability of surviving compared to individuals with a persistent fever or persistent rash. This assumption is largely based on Eramova, Matic, and Munz's (2006) analysis of general chronic symptoms.

(1) **S(t)=** Pr { β_1HIV [1-4], β_2BMI, β_3Duration of Chronic Health Symptoms, β_4Visit to Doctor, β_5Male, β_6Age > t }

(2) **S(t)=** Pr { β_1HIV [1-4], β_2BMI, β_3Duration of Chronic Health Symptoms, β_4Visit to Doctor, β_5Male, β_6Age, **β_7Education, β_8Marital Status, β_9Born in Kagera Region, β_{10}Toliet, β_{11}House with Floor, β_{12}Sex of Head of Household, β_{13}Years of Education (head of household), β_{14}Total Household Consumption** > t }

All of the above specifications contain important independent variables that estimate the differences in survival time for individuals who report chronic illnesses. For this study, the most relevant variables are the four chronic illnesses. The specific health variables (*Duration of Chronic Health Symptoms*, *Visit to Doctor, BMI)* help to explain the variation in health-related outcomes between individuals who survive with chronic symptoms and those who do not. The additional demographic variables of interest serve as control variables, but are equally important for understanding the population at risk. The design of well-targeted policy involves understanding the different demographics of the affected group. *Gender, Age, Education, Marital Status* and *Consumption* are all variables that consider some of these differences.

Despite the fact that the above specifications include a number of important independent variables, there is likely to be omitted variable bias. One variable that is not explained is a family's history of health conditions. Individuals with a shorter survival time may be more inclined to die because of a history of certain illnesses within the family. Survival time is likewise confounded by the presence of other diseases (malaria, tuberculosis and typhoid) in the Kagera region. Also, the KHDS is limited to verbal responses on issues of health and does not perform any medical exams. As a result, there is no certainty that an individual's change in survival time is a result of one chronic symptom (or any chronic symptoms). These symptoms, therefore, may not be capturing the full effects of a disease.

The verbal reporting of health issues also brings into question the accuracy of the reported symptoms. While the KHDS survey defines 'chronic symptoms' as symptoms experienced for "more than six months," it is difficult to determine how accurate an individual's understanding of chronic illnesses truly was at the time of the survey. An individual who reports experiencing multiple chronic symptoms may mistake chronic with acute symptoms.

A final limitation of this model is that it only captures the variation of independent variables at two different time periods: time of entry (1991) and time of exit (year of death, if any). Advanced survival models consider the variation of other variables at a number of different times. The model used here simplifies the survivor function and assumes away the variation in chronic symptoms between 1991 and 2004. If these symptoms really are chronic, then they should not change. However, since these symptoms are verbally reported, there is likely to be variation.

V. Results

Descriptive Statistics

In order to capture variations in survival time, this study uses a sample made up of individuals interviewed in 1991 with *wave 1* of the KHDS, and who are either re-interviewed in 2004 or recorded to have died between 1991 and 2004. Within the sample, there are 1,861 individuals aged 15 to 95, 1,373 who are reported to be alive in 2004 and 488 who are reported dead in 2004 (see Table 1). Overall, about 26 percent of the sample interviewed in 1991 is reported to have died over the following 13 years.

The sample is narrowed to include individuals between the ages of 15-95. For the purpose of capturing data on those who were married, this study disregarded individuals under the age of 15 years-old. Likewise, individuals over the age of 95 are dropped from the sample. Individuals who reported a missing value for BMI are also dropped.

The overall sample can be characterized as predominately married; female; with 7 to 9 years of schooling; and highly concentrated between the ages of 16 to 25 years old and 45 to 54 years old. Nearly 94 percent of all the individuals in the sample reported having either a toilet or latrine in their house; while only 14 percent lived in a house without a dirt floor. At the time of the 1991 survey, fewer than 50 percent of the sample reported being born in the Kagera region.

Of the entire sample, about 26 percent reported having been ill for the past six-months, that is, having a chronic illness. The most commonly reported chronic symptoms were persistent weight-loss (19 percent of the sample) and chronic fever (12 percent). Few individuals reported having more than one chronic illness (coded as HIV 1- HIV 4 in Table 1). Finally, a significant number of individuals visited a doctor.

Table 1 also disaggregates the overall sample by gender. In absolute terms, more women die than men; however, when comparing the ratio of male and female deaths to the size of each sample, a higher percentage of men die than women. There is also a small relative difference in the number of women who report chronic symptoms compared to the number of men; though, when examining the individual chronic symptoms, there is little difference between the number of men and women that record specific illnesses.

Table 1. Descriptive Statistics among individuals 15-95 years in 1991, from the Kagera Region, by gender

	All (n=1,861)		Female (n=1,035)		Male (n=826)	
Attributes	No.	(%)	No.	(%)	No.	(%)
Result After Re-contact in 2004						
Re-interviewed	1,373	(73.8)	775	(74.9)	598	(72.4)
Deceased	488	(26.2)	260	(25.1)	228	(27.6)
Chronic Symptoms						
Individuals with Chronic symptoms	488	(26.2)	271	(26.2)	217	(26.3)
Chronic Diarrhea	66	(3.5)	31	(3.0)	35	(4.2)
Persistent Fever	222	(12.0)	120	(11.6)	102	(12.3)
Rash	194	(10.4)	97	(9.4)	97	(11.7)
Persistent Weight-loss	349	(18.8)	182	(17.6)	167	(20.2)
HIV 1 (chronic fever and weight-loss)	124	(6.7)	65	(6.3)	59	(7.1)
HIV 2 (chronic diarrhea and weight-loss)	39	(2.1)	18	(1.7)	21	(2.5)
HIV 3 (rash, diarrhea, and weight-loss)	15	(0.1)	7	(0.0)	8	(0.0)
HIV 4 (chronic fever, weight-loss, rash)	31	(1.7)	16	(1.5)	15	(1.8)
Visited a doctor for Chronic Symptom	294	(15.8)	157	(15.2)	137	(16.6)
Age						
15 years-old	84	(4.5)	41	(4.0)	43	(4.2)
16-25 years-old	620	(33.3)	341	(32.9)	279	(26.9)
26-34 years-old	263	(14.1)	159	(15.4)	104	(12.6)
35-44 years-old	287	(15.4)	177	(17.1)	110	(13.3)
45-54 years-old	218	(11.7)	121	(11.7)	97	(11.7)
55-64 years-old	203	(10.9)	105	(10.1)	98	(11.9)
65-74 years-old	113	(6.1)	54	(5.2)	59	(7.1)
75+ years-old	73	(3.9)	37	(3.6)	36	(4.4)
Marital Status in 1991						
Married or Partner	968	(47.0)	492	(43.2)	476	(49.9)
Previously Married (divorced/widowed)	469	(22.6)	376	(33.0)	93	(9.9)
Never Married	636	(31.0)	270	(23.7)	366	(39.1)
Education						
0 years	498	(26.8)	389	(37.6)	109	(13.2)
1-3 years of schooling	135	(7.3)	65	(6.3)	70	(8.5)
4-6 years of schooling	461	(24.8)	196	(18.9)	265	(32.1)
7-9 years of schooling	720	(38.9)	365	(35.3)	355	(42.9)
10+ years of schooling	47	(2.5)	20	(1.9)	27	(3.3)
Read a Newspaper	1,320	(70.9)	633	(61.2)	687	(83.2)
Household Characteristics						
Toilet/latrine in house	1,757	(94.4)	970	(93.7)	787	(95.3)
Lives in house with non-earth floor	264	(14.2)	150	(14.5)	114	(13.8)
Male Head of Household	1,475	(79.2)	745	(72.9)	730	(88.4)
Born in Kagera Region	905	(48.6)	357	(34.5)	548	(66.3)

Differences among the sex-disaggregated data are most notable with regard to martial status and education. Women, in general, have fewer years of schooling then men. Nearly 40 percent of women reported having 0 years of education; while only 13 percent of men have 0 years of education. Differences in marital status are also noteworthy; 33 percent of the women in the sample reported having been previously married, and either divorced or widowed. This is significantly different from males; about 10 percent of males report being previously married. Overall, these differences prove to be significant factors with regard to survival time.

Table 2 captures the basic descriptive statistics for other variables used in this analysis. The average age of an individual from the baseline sample (1991) is about 37 years, with the oldest person being 95 years-old and the youngest person being 15 years old. Regarding education, on average, individuals in the sample have about 4 years of education; similarly, the number of years of schooling for the head of the household, on average, is about 4 years.

Annual household consumption per capita averages around 178,630.50 Tanzanian Shillings (TZS). Using the 2004 exchange rate ($1=1,130.5 TZS), this household consumption averages to about $158 USD per year. The sample shows high variation in household consumption per capita, with the lowest per capita consumption reported as 15,444 TZS (US$14) and the highest per capita consumption being 1,345,833 TZS (US$1190).

Table 2. Descriptive Statistics of Variables Used in the Analysis

Variable	Obs	Mean	Std Dev	Min	Max
General					
Gender (1=male, 0 otherwise)	1861	0.44	0.49	0	1
Age (in 1991)	1861	37.06	18.67	15	95
Result After Re-contact in 2004 (1=dead, 0 otherwise)	1861	0.26	0.44	0	1
Chronic Symptoms (=1, 0 otherwise)					
Individuals with Chronic symptoms	1861	0.26	0.44	0	1
Chronic Diarrhea	1861	0.04	0.19	0	1
Persistent Fever	1861	0.12	0.32	0	1
Rash	1861	0.10	0.31	0	1
Persistent Weight-loss	1861	0.19	0.39	0	1
HIV 1 (chronic fever and weight-loss)	1861	0.07	0.25	0	1
HIV 2 (chronic diarrhea and weight-loss)	1861	0.02	0.14	0	1
HIV 3 (rash, diarrhea, and weight-loss)	1861	0.01	0.09	0	1
HIV 4 (chronic fever, weight-loss, rash)	1861	0.02	0.13	0	1
Health Indicators					
Visited a doctor for Chronic Symptom	1861	0.16	0.36	0	1
Body Mass Index	1861	20.62	2.89	15	34
Duration of Illness (weeks)	1861	113.93	285.10	0	1560
Marital Status in 1991(=1, 0 otherwise)					
Married or Partner	1861	0.48	0.49	0	1
Previously Married (divorced/widowed)	1861	0.23	0.42	0	1
Never Married	1861	0.29	0.45	0	1
Education					
School Years	1861	4.33	3.13	0	19
Read a Newspaper (=1, 0 otherwise)	1861	0.72	0.45	0	1
Additional Characteristics					
Toilet/latrine in house (=1, 0 otherwise)	1861	0.94	0.23	0	1
Lives in house with non-earth floor (=1, 0 otherwise)	1861	0.14	0.35	0	1
Sex of Head of Household (1=male, 0 otherwise)	1861	0.79	0.41	0	1
Years of Education for Head of Household	1861	4.24	3.02	0	18
Born in Kagera Region (=1, 0 otherwise)	1861	0.49	0.50	0	1
Total Household Consumption per capita in 2004 TZS	1861	178630.50	129063.90	15444	1345833

Finally, there is significant variation in other health variables. Conditional on reporting a chronic symptom, the average individual reported having the symptom for about 114 weeks, or 2.1 years. All individuals who reported not having a chronic symptom were coded as having 0 weeks for the duration of the symptoms; while the maximum duration was reported as 30 years. BMI also ranged from 15 to 34, with an average of 20.62. This is quantified as a normal BMI (the normal range is between 18.5 and 24.9).

Multivariate Analysis

Within survival analysis, there are three main types of models: nonparametric estimation, parametric models, and Cox proportional hazards model (semi-parametric). This study employs the nonparametric estimation (or the Kaplan-Meier survival estimation) and the Cox proportional hazards model. The Kaplan-Meier survival estimation measures the probability of surviving past time t, or in this study, 1991. Nonparametric models are limited in scope, however, because they simply test the hypothesis of whether the probability of surviving for one group (e.g. males) differs from the probability of surviving in another group (e.g. females).

Cox proportional hazards models differ from both parametric and non-parametric models because they employ other variables in the multivariate analysis (e.g. time at which the record came under observation; time at which the record left the observation; a variable indicating failure (death); and a variable on whether the observation is to be used) (StataCorp, 2005). Moreover, they determine the risk of an individual surviving relative to those who are still living. The Cox model is *semi-parametric* because it assumes that the relative risk of death differs over time, in comparison with the *parametric* model, which assumes that the risk is the same over time. In this study, both the Kaplan-Meier and Cox models are used to explain the variation in survival time.

In order to analyze the relative risk of an individual surviving, the data described above were converted into survival data by establishing the two time variables (start, which is age in 1991; and end, which is age at death or age in 2004); and the failure variable (=1 if dead). After setting the data to survival analysis mode, there were 488 total failures, with the average entry age of 37 years (time 0: 1991) and the average exit age of 48 years (time 1: 2004). From 1991, the average survival time for an individual is about 11 years. Overall, 488 individuals are reported dead by 2004.

Before running a Cox proportional hazards model, the Kaplan- Meier curve was estimated for gender (see Figure 1). The curve showed no significant difference in survival probability between males and females. Results from the statistical test confirmed that there was no statistical significance for gender in survival probabilities (logrank test $p = 0.20$).

Similarly, the Kaplan- Meier Survival curve was predicted for previously married individuals (divorced, widowed, or separated) relative to individuals who have either never married or married. The curve showed that there is a significant difference between the two groups of individuals (logrank test $p=0.000$). The survival line in Figure 2 for previously married individuals shows that those who have been previously married are predicted to die sooner and at a much faster rate compared to those who never married or who are married (prevmarr=0).

Figure 1. Kaplan-Meier estimates of cumulative survival probabilities by 2004 for men and women

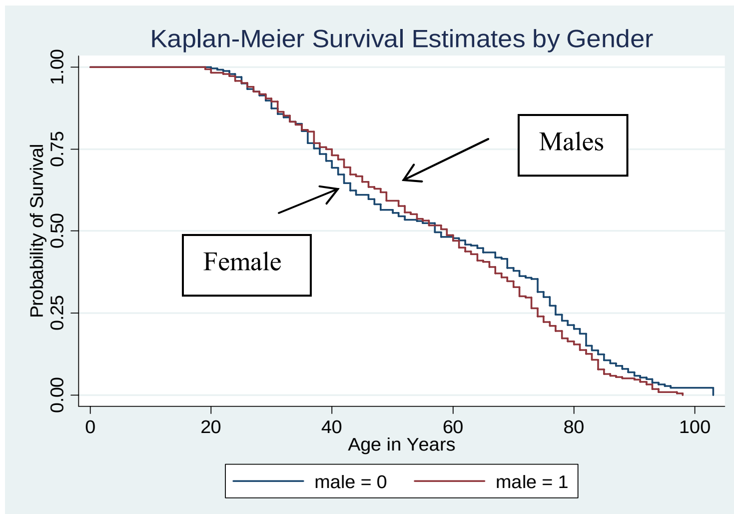
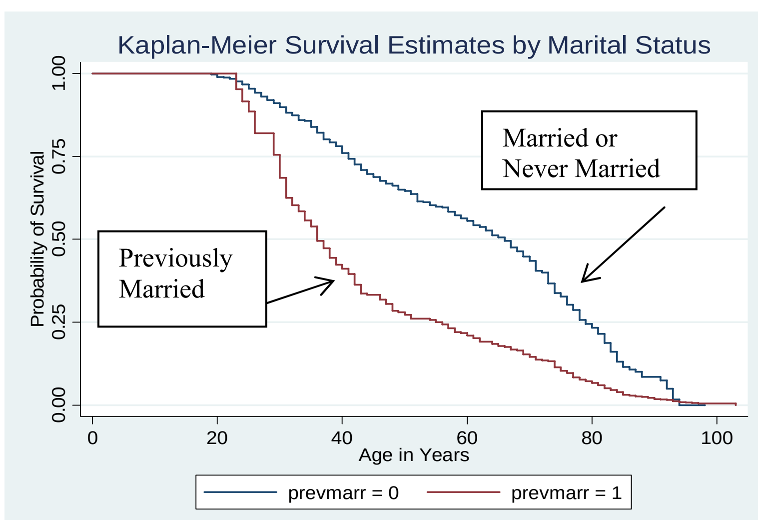


Figure 2. Kaplan-Meier estimates of cumulative survival probabilities by 2004 for individuals previously married relative baseline individuals who are either married or never married



The Kaplan-Meier model is useful for understanding the differences in survival time for individual exogenous variables. To capture the effects of multiple variables on survival time, the study also uses a Cox proportional hazards model.

Cox proportional hazards ratios for the risk of death of an individual with certain chronic symptoms are presented in Table 3 in three different specifications. In the first specification, the effect of the four chronic symptoms on survival time is evaluated. Persistent weight-loss is the only chronic symptom that is reported as statistically significant, where the risk of disease progression to death is 41 percent greater than those who do not report having chronic weight-loss. Although not statistically significant, individuals reporting the other three chronic symptoms are predicted to have a greater risk of death, relative to those individuals who do not report a chronic illness.

Controlling for other biological and demographic exogenous variables, the second and third specifications show that there are other factors that contribute to a shorter survival time. In the third specification, individuals experiencing persistent weight-loss have a 28 percent greater risk of death, compared to individuals who do not experience this chronic illness.

The variable *duration of chronic illnesses* was also statistically significant. As hypothesized, individuals who reported having a chronic illness of less than a year were likely to live longer. In the analysis, individuals who experienced a chronic illness between 53 and 104 weeks (1-2 years) presented a 64 percent greater risk of death than individuals who experience some chronic symptom between 0 and 26 weeks (6 months). Men and women who reported between 157 and 209 weeks of illness (3-4 years) have a 94 percent greater risk of death compared to the baseline group.

Visiting a doctor and receiving a medical diagnosis was shown to have no statistical significance or effect on the estimated risk of survival for individuals reporting a chronic illness compared to those who do not report an illness. While receiving a medical diagnosis for a chronic illness may suggest that an individual's survival time may be prolonged, it does not capture whether the individual received medical treatment or medicines. Furthermore, the quality of the diagnosis cannot be measured in this variable. As a result, receiving a doctor's diagnosis for a chronic illness has little effect on the estimated risk of death.

The second hypothesis in this study predicted that individuals with specific demographic characteristics had shorter survival times. Controlling for other demographic characteristics, age was only statistically significant for one group of individuals: 35-44-year-olds. Individuals aged 35-44 years old were found to have a 161 percent greater risk of progression to death compared to the baseline age group, individuals of 15-years of age. This number is alarmingly high; however, this age group is only slightly below the numbers of the national average life expectancies in Tanzania. There is a trend of increased risk of progression to death as age increases (63 percent greater risk for individuals 16-25 years old and 87 percent greater risk for individuals 26-

34 years old). Estimated hazard ratios for other age categories are not statistically significant and, therefore, are not presented in the table.

Number of years of education also affects the relative risk of survival in the sample. Surprisingly, individuals with more education are predicted to have a higher risk of death. Individuals with 4-6 years of schooling presented a 30 percent higher risk of death, than those with 0 years of schooling. Likewise, individuals with 10 or more years of education have a 130 percent greater risk of progression to death compared to individuals with 0 years of education, controlling for chronic illnesses and other biological factors. Other education variables are not presented in the table because they are not statistically significant.

As shown in the Kaplan-Meier graph above, being previously married is a predictor of shorter survival time. Those who were previously married have a 119 percent higher risk of death, compared to those who are not married. Of the 423 individuals reported as previously married, 329 are female (see Table 1). This may suggest that their husbands were carrying a chronic illness, or some other illness, and passed it along to them; however, only 116 females reported a chronic illness at the time of the KHDS in 1991. There are a number of other exogenous factors not included in this analysis that may also increase an individual's risk of survival (e.g. sexual behavior prior to marriage; sexual behavior during marriage).

Table 3. Cox Proportional Hazards Model: Estimated Risk of Death, Chronic Symptoms and other Characteristics

Variable	Multivariate Analysis: Hazard Ratios[1] 1	2	3
Chronic Symptoms (=1, 0 otherwise)			
Chronic Diarrhea	1.16 (0.24)	1.07 (0.23)	1.14 (0.24)
Persistent Fever	1.17 (0.16)	1.09 (0.15)	1.19 (0.17)
Rash	1.21 (0.17)	1.17 (0.16)	1.16 (0.17)
Persistent Weight-loss	1.41 (0.16)***	1.35 (0.16)***	1.28 (0.15)**
Body Mass Index (BMI) (baseline: BMI<18.5)			
Normal (18.5-24.9)		0.78 (0.08)**	0.78 (0.08)**
Overweight (25.0+)		0.95 (0.19)	0.90 (0.19)
Duration of Chronic Illness (weeks) (baseline: 0-26 weeks)[2]			
27-52 weeks		1.37 (0.27)	1.31 (0.26)
53-104 weeks		1.67 (0.38)**	1.64 (0.39)**
105-156 weeks		1.38 (0.32)	1.37 (0.32)
157-209 weeks		1.74 (0.45)**	1.94 (0.51)**
Illness diagnosed by doctor (=0 if chronic symptoms=0)		0.98 (0.14)	1.01 (0.15)
Individual Characteristics			
Gender (1=male, 0 otherwise)		1.12 (0.11)	1.27 (0.15)**
Age (baseline: 15 years of age)[2]			
16-25 years old		1.62 (0.77)	1.63 (0.78)
26-34 years old		2.01 (1.07)	1.87 (1.01)
35-44 years old		3.25 (1.87)**	2.61 (1.53)*
Education (baseline: 0 years of schooling)[3]			
1-3 years of education			1.24 (0.23)
4-6 years of education			1.30 (0.19)*
10+ years of education			2.30 (0.73)***
Marital Status (=1, 0 otherwise) (baseline: never married)			
Married			1.03 (0.20)

Table 3 Continued

Variable	Multivariate Analysis: Hazard Ratios[1] 1	2	3
Previously Married			2.19 (0.47)***
Additional Characteristics			
Toilet/latrine in house			0.73 (0.14)*
Lives in house with non-earth floor			1.19 (0.17)
Sex of Head of Household (1=male, 0 otherwise)			1.03 (0.15)
Years of Education, Head of Household			1.00 (0.02)
Born in Kagera Region (=1, 0 otherwise)			1.12 (0.11)
Total Household Consumption per capita in 2004 TZS (baseline: 15,444- 113,050 TZS; $14-$100 USD)			
113,501-565250 TZS ($101-$500 USD)			0.94 (0.10)
565,251-1,130,499 TZS ($501-$1000 USD)			1.02 (0.34)
1,130,500 TZS ($1001+ USD)			2.02 (1.56)
Total Number of Observations	1861	1861	1861
Total Number of Failures (dead in 2004)	488	488	488
Chi-Square Test	20.11***	54.05***	116.05***

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level; [1]These values are hazard ratios, not betas. The Cox hazard ratio is the ratio of the probability of death occurring in the observed group over the baseline group; [2] Week dummy variables were created for 210-520 weeks; and 521+ weeks; neither are statistically significant. Age dummies were created for 45-54 years; 55-64 years; 65-74 years; and 75+ years. Hazard Ratios for these variables are insignificant. [3]Dummy variables were created for 1-3 years; 4-6 years; 7-9 years and 10+ years of education. The hazard ratios for 4-6 years, and 7-9 years of education are insignificant and therefore not listed.

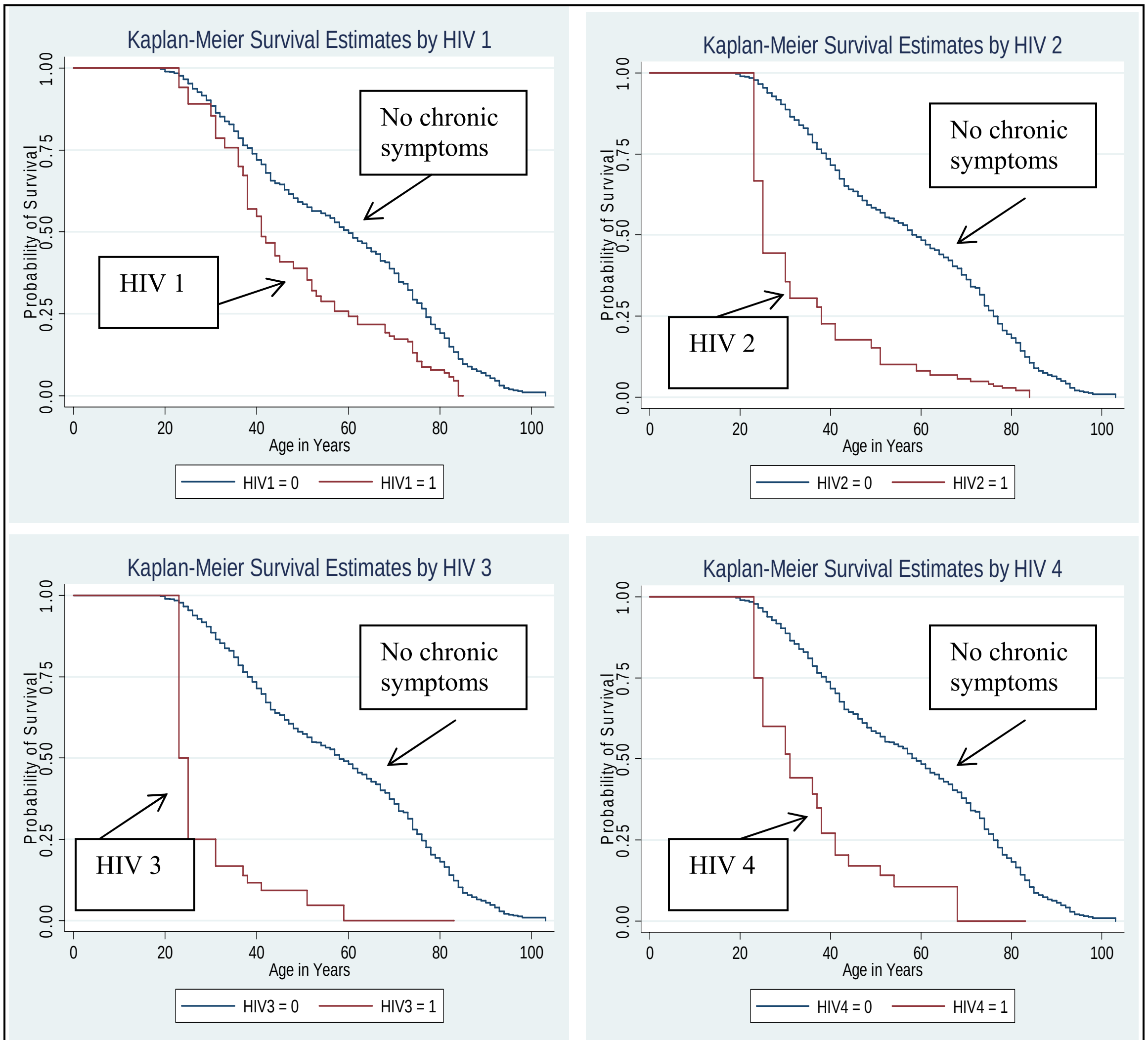
Normal body mass index (18.5-24.9) is also statistically significant. Compared to individuals who are underweight (<18.5), normal BMI is a predictor of longer survival time. Controlling for other biological factors, individuals with a normal BMI have a 22 percent lower risk of death than individuals in the baseline.

Migration, measured as whether the individual was born in the Kagera region or migrated to the Kagera region, is not statistically significant. In this study, migrants were hypothesized to have a shorter survival time because migration has been shown to increase one's risk of contracting HIV/AIDS, among other diseases (UNAIDS, 2007). However, it may be the case that migrants have better access to medical care prior to moving to the Kagera region. Even if the hazard ratio were significant, there would be little difference in the risk of death between those who were born in the Kagera region and those who were born elsewhere.

Finally, there is a difference in survival time between males and females. Although the Kaplan-Meier curve showed no statistical difference in duration of survival between males and females, when controlling for other demographic factors, females are predicted to have a shorter survival time relative to males. Females aged 15-95 in the Kagera region of Tanzania have a 27 percent greater risk of death compared to males. This is contrary to my hypothesis and to most national life-expectancy data, where males are predicted to have a higher risk of death relative to females.

Additional multivariate models were run using the generated HIV/AIDS variables. As hypothesized, individuals reporting one or more chronic symptoms, which are captured in the synthetic HIV/AIDS variables, had shorter survival times, or an increased risk of death, relative to those who did not report having a chronic illness. Kaplan-Meier curves for the four HIV/AIDS variables estimated the cumulative survival probabilities for those with multiple chronic illnesses (represented in the HIV variables) and those not living with chronic symptoms (see Figure 3).

Figure 3. Kaplan-Meier Survival Estimates for cumulative survival probabilities in 2004 for individuals with multiple chronic symptoms (constructed HIV variables)*



*Definition of generated HIV variables: HIV 1: chronic fever and weight-loss; HIV 2: chronic diarrhea and weight-loss; HIV 3: rash, diarrhea, and weight-loss; HIV 4: chronic fever, weight-loss, rash

Cox hazard models showed that all four HIV/AIDS variables reduce the likelihood of surviving. In Table 4, specifications 1 and 2 show the effects of chronic fever and weight-loss.

Controlling for both demographic and biological differences, individuals who report both chronic fever and chronic weight-loss have a 55 percent greater risk of death than individuals who do not report these chronic symptoms. In specifications 3 and 4, individuals reporting chronic diarrhea and weight-loss (HIV 2) had a 61 percent higher risk of death compared to those who did have those two chronic symptoms.

Table 4. Cox Proportional Hazards Model: Estimated Risk of Death, HIV Variables and other Characteristics

	Multivariate Analysis: Hazard Ratios[1]			
Variable	**1**	**2**	**3**	**4**
Chronic Symptoms				
HIV 1 (chronic fever and weight-loss)	1.49 (0.23)***	1.55 (0.24)***		
HIV 2 (chronic diarrhea and weight-loss)			1.60 (0.39)***	1.61 (0.39)**
Body Mass Index (BMI) (baseline: BMI<18.5)				
Normal (18.5-24.9)	0.78 (0.08)**	0.77 (0.08)**	0.76 (0.08)***	0.76 (0.08)***
Overweight (25.0+)	0.92 (0.18)	0.87 (0.18)	0.89 (0.18)	0.85 (0.17)
Duration of Chronic Illness (weeks) (baseline: 0-26 weeks)[2]				
27-52 weeks	1.46 (0.28)**	1.41 (0.27)*	1.46 (0.28)*	1.42 (0.28)*
53-104 weeks	1.74 (0.39)**	1.70 (0.40)**	1.67 (0.28)**	1.64 (0.39)**
105-156 weeks	1.48 (0.33)*	1.45 (0.33)	1.48 (0.34)*	1.44 (0.33)
157-209 weeks	1.78 (0.46)**	2.01 (0.53)***	1.75 (0.45)**	1.97 (0.52)***
Illness diagnosed by doctor (=0 if chronic symptoms=0)	0.97 (1.14)	1.01 (0.15)	0.99 (0.15)	1.04 (0.16)
Individual Characteristics				
Gender (1=male, 0 otherwise)	1.13 (0.11)	1.31 (0.16)**	1.13 (0.11)	1.31 (0.16)**

Table 4, Continued

Variable	Multivariate Analysis: Hazard Ratios[1] 1	2	3	4
Age (baseline: 15 years of age)[2]				
16-25 years old	1.68 (0.79)	1.68 (0.81)	1.71 (0.81)	1.71 (0.82)
26-34 years old	2.12 (1.12)	1.94 (1.04)	2.15 (1.14)	1.97 (1.06)
35-44 years old	3.42 (1.97)**	2.73 (1.61)*	3.47 (2.00)**	2.76 (1.63)*
Education (baseline: 0 years of schooling)[3]				
1-3 years of education		1.23 (0.23)		1.22 (0.22)
4-6 years of education		1.31 (0.19)*		1.29 (0.19)*
10+ years of education		2.25 (0.48)***		2.20 (0.70)**
Marital Status (=1, 0 otherwise) (baseline: never married)				
Married		1.04 (0.20)		1.04 (0.20)
Previously Married		2.25 (0.48)***		2.22 (0.47)***
Additional Characteristics				
Toilet/latrine in house		0.72 (0.13)*		0.71 (0.13)*
Lives in house with non-earth floor		1.17 (0.16)		1.19 (0.17)
Sex of Head of Household (1=male, 0 otherwise)		1.04 (0.15)		1.03 (0.15)
Years of Education, Head of Household		1.01 (0.02)		1.00 (0.02)
Born in Kagera Region (=1, 0 otherwise)		1.09 (0.11)		1.09 (0.11)
Total Household Consumption per capita in 2004 TZS (baseline: 15,444- 113,050 TZS; $14-$100 USD)				
113,501-565250 TZS ($101-$500 USD)		0.95 (0.10)		0.93 (0.10)
565,251-1,130,499 TZS ($501-$1000 USD)		1.03 (0.34)		1.04 (0.34)
1,130,500 TZS ($1001+ USD)		2.30 (1.77)		2.31 (1.78)
Total Number of Observations	1861	1861	1861	1861
Total Number of Failures (dead in 2004)	488	488	488	488
Chi-Square Test	48.31***	112.05***	45.86***	108.23***

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level; [1]These values are hazard ratios, not betas. The Cox hazard ratio is the ratio of the probability of death occurring in the observed group over the baseline group; [2] Dummy variables created for 210-520 weeks; and 521+ weeks; neither are statistically significant. Age dummies created for 45-54 years; 55-64 years; 65-74 years; and 75+ years. Hazard Ratios are insignificant. [3]Dummy variables created for years of education. Some hazard ratios are insignificant and therefore not listed.

Similarly, Table 5 shows the effects on survival for individuals who report at least three chronic symptoms, defined as HIV 3 and HIV 4. Controlling for other biological and demographic variables, individuals who report having HIV 3 (rash, diarrhea, and chronic weight-loss) have a 126 percent greater risk of death compared to those who do not report these chronic symptoms.

Table 5. Cox Proportional Hazards Model: Estimated Risk of Death, HIV Variables and other Characteristics

	Multivariate Analysis: Hazard Ratios[1]			
Variable	**1**	**2**	**3**	**4**
Chronic Symptoms				
HIV 3 (rash, diarrhea, and weight-loss)	2.34 (0.83)**	2.26 (0.82)**		
HIV 4 (chronic fever, weight-loss, rash)			2.19 (0.58)***	2.38 (0.64)***
Body Mass Index (BMI) (baseline: BMI<18.5)				
Normal (18.5-24.9)	0.75 (0.08)***	0.74 (0.08)***	0.75 (0.08)**	0.74 (0.08)***
Overweight (25.0+)	0.87 (0.17)	0.83 (0.17)	0.88 (0.17)	0.83 (0.17)
Duration of Chronic Illness (weeks) (baseline: 0-26 weeks)[2]				
27-52 weeks	1.45 (0.28)*	1,42 (0.28)*	1.45 (0.28)*	1.39 (0.27)*
53-104 weeks	1,73 (0.39)**	1.72 (0.40)**	1.76 (0.40)**	1.74 (0.41)**
105-156 weeks	1.53 (0.35)*	1.49 (0.34)*	1.50 (0.34)*	1.45 (0.33)
157-209 weeks	1.76 (0.45)**	1.97 (0.52)***	2.78 (0.46)**	1.98 (0.52)***
Illness diagnosed by doctor (=0 if chronic symptoms=0)	0.98 (0.14)	1.02 (0.15)	0.97 (0.14)	1.02 (0.15)
Individual Characteristics				
Gender (1=male, 0 otherwise)	1,12 (0.11)	1.31 (0.16)**	1.12 (0.11)	1.31 (0.16)**
Age (baseline: 15 years of age)[2]				
16-25 years old	1.74 (0.82)	1.73 (0.83)	1.72 (0.81)	1.72 (0.82)
26-34 years old	2.19 (1.16)	2.00 (1.08)	2.18 (1.15)	1.98 (1.07)
35-44 years old	3.52 (2.02)**	2.78 (1.64)*	3.50 (2.02)**	2.76 (1.63)*

Table 5, Continued

Variable	Multivariate Analysis: Hazard Ratios[1] 1	2	3	4
Education (baseline: 0 years of schooling)[3]				
1-3 years of education		1.20 (0.22)		1.23 (0.23)
4-6 years of education		1.20 (0.19)*		1.32 (0.19)*
10+ years of education		2.20 (0.70)**		2.27 (0.72)**
Marital Status (=1, 0 otherwise) (baseline: never married)				
Married		1.04 (0.20)		1.05 (0.20)
Previously Married		2.20 (0.47)***		2.26 (0.48)***
Additional Characteristics				
Toilet/latrine in house		0.70 (0.13)*		0.71 (0.13)*
Lives in house with non-earth floor		1.18 (0.16)		1.19 (0.17)
Sex of Head of Household (1=male, 0 otherwise)		1.01 (0.14)		1.02 (0.15)
Years of Education, Head of Household		1.00 (0.01)		1.00 (0.02)
Born in Kagera Region (=1, 0 otherwise)		1.09 (0.11)		1.09 (0.11)
Total Household Consumption per capita in 2004 TZS (baseline: 15,444- 113,050 TZS; $14-$100 USD)				
113,501-565250 TZS ($101-$500 USD)		0.93 (0.10)		0.93 (0.09)
565,251-1,130,499 TZS ($501-$1000 USD)		1.04 (0.34)		1.04 (0.34)
1,130,500 TZS ($1001+ USD)		2.21 (1.70)		2.17 (1.67)
Total Number of Observations	1861	1861	1861	1861
Total Number of Failures (dead in 2004)	488	488	488	488
Chi-Square Test	49.90***	108.81***	49.59***	113.10***

* Significant at the 10% level; ** Significant at the 5% level; *** Significant at the 1% level; [1]These values are hazard ratios, not betas. The Cox hazard ratio is the ratio of the probability of death occurring in the observed group over the baseline group; [2] Dummy variables created for 210-520 weeks; and 521+ weeks; neither are statistically significant. Age dummies created for 45-54 years; 55-64 years; 65-74 years; and 75+ years. Hazard Ratios are insignificant. [3]Dummy variables created for years of education. Some hazard ratios are insignificant and therefore not listed.

Individuals who reported having chronic fever, weight-loss, and rash (coded as HIV 4) are estimated to have a shorter survival time; individuals with chronic symptoms of HIV 4 have a 138 percent greater risk of death than individuals without these symptoms. This affirms the hypothesis that individuals with more than one chronic illness are predicted to have a greater risk of death, relative to those who do not report multiple symptoms.

VI. Discussion

While the use of the KHDS panel data containing demographic and biological information allows for the analysis of survival time, these data have many limitations. First, the variables used to capture the chronic health symptoms are very limited in scope. There is no medical examination or in-depth analysis of each symptom. Further, these data rely heavily on the individual's ability to consistently self-report the chronic symptoms they have been experiencing and for how long. If the data are reported inconsistently, the results in this analysis are likely to be biased.

Second, in the multivariate analysis, of the four chronic symptoms evaluated in this study, chronic weight-loss alone was statistically significant and the strongest predictor of risk of death. There are a number of exogenous factors that could be biasing the hazard ratio on the chronic weight-loss variable. Some of these factors include: poor nutrition, acute or chronic illness, or a disease inherent to the family. To control for poor nutrition, the independent variable *BMI* was included in the model. To ensure there was no endogeneity bias, *Pearson's r* was calculated for BMI and weight-loss; there is a weak-negative correlation ($r = -0.1414$).

Other exogenous factors such as history of family illnesses or other chronic illnesses (e.g. chronic cough; frequent nausea; chest pain) may also bias the hazard ratio (risk of death) for persistent weight-loss or any of the other chronic symptoms. This dataset is limited to the four HIV/AIDS symptoms. Including other chronic symptoms in future analyses may help explain which chronic symptoms are most deadly.

The generated HIV variables also have a number of limitations. Logical health patterns suggest that an individual with multiple chronic health symptoms is likely to die faster than an individual with no chronic illnesses, as observed in this study's analysis. While all of the generated HIV variables are statistically significant, they too are likely to be biased. If there are exogenous variables that affect any of the four chronic symptoms, these exogenous factors will also bias the four generated HIV variables. Either way, all four chronic symptoms predict a higher risk of death as hypothesized.

The predicted effect of the number of school years on the estimated risk of survival also seems somewhat counterintuitive and contrary to the hypothesis in this study. In the multivariate analysis, the marginal effect of an additional 4-6 years of education or 10-plus years of education results in a higher risk of death, relative to those who have no education. In a recent demography study in the U.S. relating health and education, Cutler and Lleras-Muney (2007) found that better educated people have lower death rates resulting from common acute illnesses and chronic diseases. While this may be the case for U.S. health and education, the causal relationship between education and

HIV/AIDS in the Kagera region in Tanzania has shown to be somewhat less than straightforward.

In contrast to Cutler and Lleras-Muney, research completed in the early 1990s suggests that higher levels of education are correlated with higher HIV prevalence (World Food Programme, 2004). More recent evidence, however, suggests that there are likely to be other exogenous factors that lead to this positive correlation e.g., gender inequalities, lack of access to work, and income inequalities (Over, 2001). While the model controls for gender and annual household consumption, the outcome may suggest that the number of years of schooling is a poor measure of how much knowledge individuals have about health issues.

Omitted variables that may be influencing the hazard ratio on *education* include: quality of the education; accuracy in reporting years of education; ability to read and write; and knowledge of specific health indicators. In order for the validity of these hazard ratios to be strengthened, future analyses need to include additional information about an individual's knowledge of health issues.

Another somewhat surprising finding from this study is the decreased likelihood of survival for individuals who have been previously married, relative to those who have never been married. Higher risk of death among the previously married could be caused by a number of factors. First, the previously married single male or female may not have access to medical care or attention if they are sick. Second, those who are previously married are more vulnerable to poverty because they are less likely to have access to income-smoothing mechanisms. Being married has been shown to significantly increase employment levels, incomes, and remittances for those in rural areas (Luke and Munshi, 2003). Third, it is possible that those who are previously married are widows of HIV/AIDS victims. If this is true, then the probability of surviving will decrease dramatically.

Finally, the difference in survival time for males and females is striking. When additional demographic variables are included in the multivariate model, females are predicted to have a shorter survival time than males, even though more men die in the sample. While the difference is relatively small (28 percent higher risk of death), there may be a number of exogenous factors that influence this differential. Females, for example, may have limited access to health care relative to males. With limited access to health care, women are more vulnerable disease progression.

The literature on HIV/AIDS explains the varied incidence of this disease by differences in gender and income inequalities. Disproportionate infection rates are largely attributed to gender-based risks and vulnerabilities such as social-cultural norms, gender-based violence, denial of rights to property or inheritance, and lack of access to economic opportunities. HIV/AIDS is fueled by key economic, socio-cultural, legal, and physiological factors that are different for women and men (World Bank, 2007). More research is needed in this region to understand why women are less likely to survive relative to males.

Overall, the present analysis is limited in its conclusions about the cause of death. While a combination of the chronic variables predicts shorter survival time, there can be no inference made about the certainty of whether these individuals have HIV/AIDS. What can be said, however, is that individuals reporting persistent weight-loss or more than one of the chronic illnesses in the Kagera region of Tanzania are predicted to have a greater risk of death relative to those who do not report chronic symptoms.

Using panel data from the Kagera region in Tanzania, the present study shows that there are a number of vulnerable populations in the Kagera region: individuals with persistent weight-loss; females; individuals aged 35-44 years of age; previously married men and women; and individuals with higher levels of education.

From a policy perspective, knowledge of the most vulnerable groups is important for preventative health care delivery. With scarce resources, however, health care providers and policymakers must also decide how to distribute resources for effective service delivery. Future cost-benefit analysis is needed to determine policies how best to use Tanzania's health resources and which populations to target.

VII. Bibliography

Ainsworth, Martha and Innocent Semali. 1998. “Who is most likely to die of AIDS? Socioeconomic Correlates of Adult Deaths in Kagera Region, Tanzania.” In *Confronting AIDS: Evidence for Developing World,* edited by M. Ainsworth, L. Fransen and M. Over. Brussels: European Commission and The World Bank.

Ainsworth, Martha and Julia Dayton. 2001. “The Impact of the AIDS Epidemic on the Health of the Elderly in Tanzania.” Washington, D.C.: The World Bank.

Ainsworth, Martha, Kathleen Beegle, Paurvi Bhatt, Vajeera Dorabawila, Daniel Dorsainvil, Indrani Gupta, Mead Over, and Jim Shafer. December 2004. “User’s Guide to the Kagera Health and Development Survey Datasets.” Washington, D.C.: The World Bank.

Aral, Sevgi O., Mead Over, Lisa Manhart and King K. Holmes. 2006. "Sexually Transmitted Infections." In *Disease Control Priorities in Developing Countries (2nd Edition),* Chapter 17, 311-330. New York: Oxford University Press.

Bagdonavicius, Vilijandas and Mikhail Nikulin. 2004. “Statistical Modeling in Survival Analysis and its Influence on the Duration of Analysis.” In *Handbook of*

Statistics: *Advances in Survival Analysis, Volume 23,* edited by N. Balakrishnan and C. R. Rao. Amsterdam: Elsevier North-Holland.

Beegle, Kathleen, Joachim De Weerdt, and Stefan Dercon. 2006a. "Adult Mortality and Consumption Growth in the Age of HIV/AIDS." Washington, D.C.: The World Bank.

______. 2006b. "Kagera Health and Development Survey 2004: Basic Information Document." Washington, D.C.: The World Bank.

Belgian Development Cooperation. 2004. "Belgian Technical Cooperation in Tanzania." Accessed on January 11, 2008 at http://www.btcctb.org/doc/UPL_200409291730173906.pdf

Boerma, J. Ties, Juliana Ngalula, Raphael Isingo, Mark Urassa, Kesheni Senkoro, R. Gabone, and E.N. Makumbo. 1997 "Levels and Causes of Adult Mortality in Rural Tanzania with Special Reference to HIV/AIDS." *Health Transition Review*: Vol. 7, Supplement 2, pp.63-74.

Braga, Patricia, Maria-Regina Cardoso, Aluisio Segurado. 2007 "Gender Differences in

Survival in an HIV/AIDS Cohort from São Paulo, Brazil." *AIDS Patient Care and STDs:* Vol. 21, No. 5. pp. 321-328.

Chapoto, Antony and T.S. Jayne. 2005. "Characteristics of Individuals Afflicted by AIDS- Related Mortality in Zambia." Food Security Research Project. Working Paper No. 14. Lusaka, Zambia: Michigan State University.

Crampin, Amelia, Sian Floyd, Judith Glynn, Felix Sibande, Dominic Mulawa, Andrew Nyondo, Philip Broadbent, Lyn Bliss, Bagrey Ngwira, and Paul Fine. 2002. "Long-Term Follow-up of HIV-Positive and HIV-Negative Individuals in Rural Malawi." *AIDS*: Vol. 16: 1545-1550.

Crush, Jonathan, Bruce Frayne, and Miriam Grant. 2006. "Linking Migration, HIV/AIDS and Urban Food Security in Southern and Eastern Africa." Washington, D.C.: International Food Policy Research Institute.

Cutler, David and Adriana Lleras-Muney. 2007. "Education and Health." National Poverty Center: Policy Brief No. 9. Ann Arbor, Michigan: Gerald R. Ford School of Public Policy, University of Michigan.

Doctor, Henry and Alexander Weinreb. 2003a. “Estimation of AIDS Adult Mortality by Verbal Autopsy in Rural Malawi.” *AIDS:* Vol. 17. pp. 2509-2513.

______ . 2003b. “Mortality among Married Men in Rural Kenya and Malawi: A Life Table Analysis.” Population Studies Center. Philadelphia, PA: University of Pennsylvania.

Eramova, Irina, Srdan Matic, Monizque Munz. 2006. “Chapter 2: Management of Opportunistic Infections and General Symptoms of HIV/AIDS.” In *HIV/AIDS Treatment and Care: Clinical Protocols for the WHO European Region.* Copenhagen, Denmark; World Health Organization, Europe:

Gutierrez, Roberto G. 2007. “Survival Analysis Using Stata.” 2007 Seminars on Stata. PowerPoint Presentation in Washington, D.C. on November 2, 2007.

Jenkins, Stephen. 2005. “Survival Analysis.” Institute for Social and Economic Research. United Kingdom: University of Essex.

Kamali, Anatoli, Hans-Ulrich Wagner, Jessica Nakiyingi, Imelda Sabiiti, Jane Kengeya-Kayondo, and Daan Mulder. 1996. “Verbal Autopsy as a Tool for Diagnosing

HIV- Related Adult Deaths in Rural Uganda." *International Journal of Epidemiology:* Vol. 25, No. 3. pp. 679-684.

Kennedy, Peter. 2003. *A Guide to Econometrics, Fifth Edition.* Cambridge, Massachusetts: MIT Press.

Killewo, J Z J., Sandstrom, A., Bredberg Raden, U., Mhalu, F.S., Biberfeld, G., and Wall, S. 1993. "Incidence of HIV-1 Infection among Adults in the Kagera Region of Tanzania." *International Journal of Epidemiology.* Vol. 22, No. 3. pp. 528-536.

Lopez, Alan D., Stephen Begg, and Ed Bos. 2006. "Demographic and Epidemiological Characteristics of Major Regions, 1990—2001." *Global Burden of Disease and Risk Factors,* Chapter 2: 17-44. New York: Oxford University Press.

Luke, Nancy and Kaivan Munshi. 2003. "New Roles for Marriage in Urban Africa." Providence, RI: Population Studies and Training Center, Brown University.

Lundberg, Mattias, Mead Over, and Phare Mujunja. 2003. "Do Savings Predict Death? Precautionary Savings during an Epidemic." Washington, D.C.: The World Bank.

Mathers, Colin D., Alan D. Lopez, and Christopher J. L. Murray. 2006. "The Burden of Disease and Mortality by Condition: Data, Methods, and Results for 2001." 2006. *Global Burden of Disease and Risk Factors,* Chapter 3: 45-93. New York: Oxford University Press.

McGready, John. 2006. “Regression for Survival Analysis.” *Johns Hopkins Bloomberg School of Public Health.* PowerPoint Presentation. Washington, D.C.: The Johns Hopkins University: Internet. Accessed on January 18, 2008 at http://ocw.jhsph.edu/courses/StatisticalReasoning2/PDFs/Lecture10.pdf.

Miller, Brian and J. Hurley.2003. “Life Table Methods for Quantitative Impact Assessments in Chronic Mortality.” *Journal of Epidemiology and Community Health*. BMJ Publishing Group. 57:200-206.

Mshana, M. Ndyetabura, E.F., Nhalu, F. Swai, R.O. 1998. “HIV/AIDS/STIs and Gender Issues in Tanzania.” International Conference AIDS: June 28-July 3, 1998. Abstract no. 14216. National AIDS Control Programme: Dares- Salaam, Tanzania. Internet. Accessed on January 12, 2008 at http://gateway.nlm.nih.gov/MeetingAbstracts/ma?f=102228083.html

Over, Mead. 2001. "The Effects of Societal variables on urban rates of HIV Infection in Developing Countries: An Explanatory Analysis." In *Confronting AIDS: Evidence for Developing World,* edited by M. Ainsworth, L. Fransen and M. Over. Brussels: European Commission and The World Bank.

StataCorp. 2005. *Stata Statistical Software: Release 9.* College Station, TX: StatCorp LP.

Sinha, Nistha and Natsujko Kiso. 2007. "HIV/AIDS Prevention and Youth: Changing Knowledge, Attitudes and Behavior." *Development Outreach.* Washington, D.C.: The World Bank.

Reed, Geoffrey, Margaret Kemeny, Shelley Taylor, Hui-Ying Wang, and Barbara Visscher. 1994. "Realistic Acceptance as a Predictor of Decreased Survival Time in Gay Men with AIDS." *Health Psychology*: Vol. 13, No.4, pp 299-307.

Rodríguez, Germán. 2007. "Chapter 7: Survival Models." Lecture Notes, WWS509. New Jersey: Princeton University. Accessed on November 1, 2007 at http://data.princeton.edu/wws509/notes/default.htm.

UCLA Academic Technology Services. 2008. "Introduction to Survival Analysis with

STATA Seminar." Los Angeles, California: UCLA. Accessed on January 15, 2008 at http://www.ats.ucla.edu/STAT/stata/seminars/stata_survival/default.htm.

UNAIDS. 2007. "AIDS Epidemic Report, 2007." Geneva, Switzerland: United Nations Development Program.

United Nations Population Fund. 2005. "Profile on Tanzania." Geneva, Switzerland: UNFPA. Accessed on January 11, 2008. http://www.unfpa.org/profile/tanzania.cfm.

Wooldridge, Jeffery M. 2006. *Introductory Econometrics: A Modern Approach,* Third Edition. Mason, Ohio: Thomson South-Western.

World Bank. 2007. "HIV/AIDS and Gender Equity." Gender and Development Briefing Note. Washington, D.C.: The Poverty Reduction and Economic Management Group, Gender Anchor. Accessed in August 2007. www.worldbank.org/gender.

World Bank. 2001. Engendering Development. Washington, D.C.: The World Bank Group.

World Food Programme. 2004. “Literature Review on the Impact of Education Levels on HIV/AIDS Prevention Rates.” Rome, Italy: World Food Programme.

World Health Organization. 2006. “Country Health System Fact Sheet 2006, United Republic of Tanzania.” Internet. Accessed on November 20, 2007. http://www.afro.who.int/home/countries/fact_sheets/tanzania.pdf.

______. February 2005. “Health Action in Crisis: Tanzania.” Internet. Accessed on November 20, 2007. http://www.who.int/hac/crises/tza/background/Tanzania_Feb05.pdf.

Yaffee, Robert. 2001. “Survival Analysis with STATA.” ITS Academic Computing Services. PowerPoint presentation. New York: New York University. Accessed on January 7, 2008 at http://www.nyu.edu/its/statistics/Docs/Survival8.ppt.

Zangerle, Robert, Gilbert Reinbnegger, and Jean-Paul Klein. 1995. “Survival Differences in Austrian Patients with the Acquired Immunodeficiency Syndrome.” *European Journal of Epidemiology*: Vol. 11, No. 5. pp. 519-526.

Zhou, Jialun and N. Kumarasamy. 2005. “Predicting short-term disease progression among HIV-infected patients in Asia and the Pacific Region: preliminary results from the TREAT Asia HIV Observational Database (TAHOD)." *HIV Medicine.* British HIV Association: Vol. 6. pp. 216-223.