Determining predictors of mortality in HIV positive people in South Africa, 2003 to 2009: A mixed methods approach incorporating unobserved variables

by

Kennedy S. Naviava Otwombe

Supervisors

1. Prof. Tobias Chirwa
2. Prof. Max Petzold
3. Dr. Neil Martinson

A THESIS

Submitted to the School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, in fulfilment of the requirements for the degree

Of

Doctor of Philosophy

Johannesburg

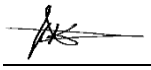
South Africa

02 April 2018

Declaration

I, Kennedy S. Naviava Otwombe declare that this thesis is my original work. Where there has been contribution from other people, this has been duly acknowledged. It is being submitted for the degree of Doctor of Philosophy in the School of Public Health, University of the Witwatersrand, Johannesburg, South Africa. It has not been submitted before for any degree or examination at this or any other University.

Name: Kennedy S. Naviava Otwombe

Signature:  Date: 02 April 2018

Dedication

I dedicate my PhD research work to my supportive family that was always there for me as I worked through my research. My dear wife Lucy and our two children Liam and Ivanna have had to endure my long periods of absence in pursuit of this degree. The support of my parents in encouraging me to complete this PhD gave me further momentum. It is because of your love and support that I dedicate this PhD to you all.

Original Papers

1. **Otwombe KN**, Petzold M, Martinson N and Chirwa T (2014). A review of the study designs and statistical methods used in the determination of predictors of all-cause mortality in HIV-infected cohorts: 2002-2011. *PLoS ONE* 9(2): e87356. doi:10.1371/journal.pone.0087356.

   Student's contribution to the paper

   Conceptualisation of the idea, collection of data, analyses of data and writing up the manuscript.

2. **Otwombe KN**, Petzold M, Martinson N and Chirwa T. Modeling clustered semiparametric frailty data: a comparison of methods using simulations (Re-submitted to *BMC Medical Research Methodology-July 2017*).

   Student's contribution to the paper

   Conceptualisation of the idea, modelling, running simulations, analyses of all simulated and actual data and writing up the manuscript.

3. **Otwombe KN**, Petzold M, Modisenyane T, Martinson N and Chirwa T (2014). Factors associated with mortality in HIV-infected people in rural and urban South Africa. *Global Health Action*, 7: 25488.

   Student's contribution to the paper

   Conceptualisation of the idea, analyses of actual study data and writing up the manuscript.

4. **Otwombe KN**, Petzold M, Martinson N, Chirwa T, Dietrich J (*To be submitted*). Understanding perceived risk factors for mortality from friends and relatives of demised HIV-infected people in Soweto, South Africa: A qualitative study.

Conference Proceedings

1. **Otwombe KN**, Petzold M, Modisenyane T, Martinson N and Chirwa T (2014). Predictors of all-cause mortality in HIV-infected people in South Africa: a comparison between rural and urban South Africa. Poster presentation at the University of the Witwatersrand, Faculty of Health Sciences, Biennial Research Day and Postgraduate Expo, 17th September 2014, Johannesburg, South Africa.

Awards

1. Doctoral fellow of the Consortium for Advanced Research Training in Africa (CARTA): 2011-2014.
2. Consortium for Advanced Research Training in Africa (CARTA) Visiting fellowship for completion of Doctoral Thesis at Centre for Applied Biostatistics, Sahlgrenska Academy, University of Gothenburg April 7 to July 2015.
3. Canada-Africa Prevention Trials (CAPT) network doctoral scholarship 2012-2013.

Abstract

Background

The largest proportion of HIV-infected people resides in Southern Africa. In South Africa, the government has taken the lead in the provision of free HIV treatment with a high coverage rate. Provision of free antiretroviral treatment has led to a decline in mortality rates and an increase in life expectancy. However, a significant number of people with HIV continue to die despite the availability of free treatment. A large proportion of studies have concentrated on using quantitative methods of analysis. Very few have used mixed methods that combine quantitative time-to-event frailty models and qualitative methods in assessing risk factors for mortality in HIV-infected individuals. However, use of such mixed methods approach could provide insights that may lead to an improvement in patient care and management.

Aim

To determine mortality risk factors in HIV-infected people through incorporating unobserved variables using a mixed methods approach in which quantitative findings are explained by the qualitative.

Methods

To critically review statistical methods used for assessing risk factors for mortality in HIV-infected people between the years 2002 and 2011. We conducted a literature review on the design of studies, how data were analysed and whether suitable statistical methods were utilised in assessing mortality risk factors in HIV-infected people in the period 2002-2011. Only publications written in English and listed in Pubmed/Medline were considered. In this review, papers using time-to-event techniques were regarded as appropriate. Data were split into two equal periods allowing for the comparison of the statistical methods over time.

To compare the different time-to-event methods, we ran 1,000 simulations of parametric clustered data using parameters derived from an HIV study that was conducted in South Africa by the Perinatal HIV Research Unit (PHRU). Data for 5, 10 and 20 clusters of size 50 and 100 were simulated. Survival and censoring times were derived from a Weibull distribution. The minimum of survival and censoring times was taken as the study time. Using the simulated data, we compared the following time-to-event methods: Cox proportional hazards regression, shared Gamma frailty with Weibull and exponential baseline hazards (frequentist models), and the Bayesian integrated nested Laplace approximation (INLA) with Weibull baseline hazard. Parameter estimates, standard errors and their fit statistics were averaged over 1,000 simulations. Similarly, means and standard deviations from INLA were averaged (over the 1,000 simulations). Frequentist models were compared using the -2 loglikelihood fit statistics while all the four models were compared using the mean square error (MSE). Additionally, we simulated semi-parametric clustered frailty models (using gamma and log-normal frailties) including INLA, h-likelihood, penalized likelihood and penalised partial likelihood estimations. Parameter estimates and their standard errors were presented graphically and compared using the MSE.

To assess mortality risk factors in HIV-infected people in South Africa in different settings, factors associated with mortality in HIV-infected people were assessed by INLA survival frailty model using cohort data of HIV-infected people from South Africa. Two thirds were from Soweto (urban) and the rest from Mpumalanga (rural). Findings were evaluated by site.

Mixed methods were used to evaluate risk factors for mortality by combining the best fitting model applied to retrospective data and qualitative analysis on prospective

data. In order to explain the unobserved frailty modelling results, we conducted a qualitative study that enrolled 20 participants who had confirmed knowing a person that had died as a result of HIV. Participants were recruited from the Zazi VCT in PHRU and were interviewed using a semi-structured interview guide. The aim of the qualitative study was to attempt to explain the unobserved factors influencing mortality in HIV-infected individuals using perceived reasons for death given by the participants. These were later used to complement the potential reasons for death as identified in the frailty modelling (quantitative) results.

Results

In the critical review, 189 studies met the inclusion criteria that included prospective (69%) and retrospective (30%) studies. Of the 189 studies, 91 were published in the period 2002-2006 and 98 in 2007-2011. Cox regression analysis with frailty was used in only 7 studies (~4%); of which 6 were published between the years 2007-2011.

The simulation study showed that the shared frailty models performed better than Cox-PH. Within the shared frailty models, the Gamma frailty model with a Weibull baseline performed better than the Gamma frailty model with an exponential baseline. The MSE showed that in general, the Bayesian INLA had better results. In the semiparametric simulations, results were similar but INLA had a slightly better fit with consistently lower MSE values relative to both gamma and log-normal frailty models. The random effects estimate for INLA, whose method is slightly different, had lower MSE values consistently relative to the other methods.

In the HIV cohort study, 6,690 participants were enrolled with majority being female (78%) and most participants residing in an urban area (67%). Rural participants were older (36 years; IQR: 31-44) and with a higher mortality rate (11/100 person years).

Among those residing in rural areas, HAART treatment for between six and twelve months (HR: 0.2, 95% CI: 0.1-0.4) and more than 12 months (HR: 0.1, 95% CI: 0.1-0.2) was protective relative to not being on treatment. Being on HAART treatment for greater than twelve months was protective in the urban participants (HR: 0.35, 95%CI: 0.27-0.46). Significant heterogeneity, assessed by frailty variance, was high in rural participants and lower in the urban.

Since the frailty modelling results suggested that the unobserved variables had a significant effect on mortality in HIV-infected individuals, a qualitative study was conducted to explore the potential causes of death. In the qualitative study, participants perceived that mortality in HIV-infected individuals may have been influenced by engagement in risky sexual behaviour such as multiple sexual partnerships, negative attitude by healthcare workers towards HIV-infected people, believing in the healing power of religion, traditional medicine, food security and social support structure.

Conclusions

The study found that Cox proportional hazards regression with frailty is not commonly used in research on mortality in HIV-infected individuals as it is used in other fields of health research. Additionally, use of the more complex semiparametric frailty models was even lower in this population. From simulations, we found that frailty survival models provided a better fit in modelling mortality due to their ability to account for unobserved variables especially the Bayesian INLA. As the unobserved variables are complex to explain using only quantitative modelling techniques, qualitative analysis of perceived causes of death was explored. Unobserved variables affecting mortality were explored through qualitative analysis of perceived reasons provided by bereaved participants. This mixed methods approach optimised

data by using a quantitative approach followed by a qualitative one that complemented each other. Use of optimal methods in assessing morbidity and mortality in HIV-infected patients may improve patient care and management in South Africa and other countries.

Key words: HIV, Mortality, Rural, Urban, unmeasured variables, HAART, Frailty

Acknowledgements

The success of this work is owed to all my supervisors. Given the intertwined aspect of moving from quantitative to qualitative analysis, I required a variety of input from various people with different expertise. On the quantitative side, I greatly thank my supervisors Prof Tobias Chirwa, Prof Max Petzold and Dr Neil Martinson for their input. Prof Tobias Chirwa's input was extremely helpful from the protocol development stage, guiding me into shaping the PhD research objectives and fitting statistical models. I will never forget the times we ran simulations for weeks with Prof Max Petzold, Health Metrics Department, University of Gothenburg, Sahlgrenska Academy, Sweden. Dr Neil Martinson, the current CEO of Perinatal HIV Research Unit (PHRU), provided clinical, epidemiological and funding support. Dr Neil's input was very helpful towards addressing the epidemiological implications of the research findings. The development of the protocol for the qualitative study, its implementation, analysis and interpretation was supported by Dr Janan Dietrich of PHRU. PHRU as an organisation was very patient and granted me time off to complete the doctoral research program. Others within PHRU were supportive in many ways that encouraged me to keep going. They include Dr Avy Violari, Prof Glenda Gray, Dr Fatima Laher and Tracey Heritage.

I will never forget the chance meeting I had with Dr Thomas Achia where he introduced me to INLA, a model that I had never worked with but ended up forming a critical part of my doctoral work. Thank you Thomas for your insights that enabled me finalise the modelling process. From you all, I have learnt not only the academic aspect of work but the importance and benefit of working as a team. This work would never have succeeded without the cooperation and collaboration that was provided by you.

I received funding from a variety of sources that enabled me to focus on my PhD. I would like to thank Consortium for Advanced Training in Africa (CARTA) for providing me with a fellowship that supported me during the entire period of study. The opportunity to take a month off once a year was very beneficial in meeting PhD targets. I am also grateful for the fellowship support to visit the Health Metrics Department, Sahlgrenska Academy, University of Gothenburg for three months to complete my dissertation. Dr Neil Martinson provided an academic scholarship in my first year through the South African Tuberculosis AIDS Training (SATBAT) grant that was funded by the Carnegie Foundation. I will forever be grateful to my employers the Wits Health Consortium through the Perinatal HIV Research Unit (PHRU) for gladly allowing me to take time off work and focus on my PhD.

Many friends and colleagues at the University of the Witwatersrand were equally encouraging and urged me on. They include Dr Anthony Essien, Dr Eustasius Musenge, Dr Adiel Chikobvu and Dr Steve Mphatso, a CARTA fellow in my cohort.

The School of Public Health at the University of the Witwatersrand was instrumental in providing me with support that enabled me to build capacity especially in methodological issues. The many seminars and workshops we attended on writing an integrated thesis facilitated by Prof Kathy Kahn were of great help.

Finally the support of fellow postgraduate students was quite encouraging. The times we spent together in the postgraduate hub at the School of Public Health were instrumental in keeping the fire burning.

This work would never have succeeded without the participants that were recruited into the Adult Wellness Study and the subsequent qualitative study that was done in PHRU. We owe the success of this study to you all.

Table of contents

List of Tables

List of Figures

List of abbreviations/acronyms

AIDS	Acquired Immune Deficiency Syndrome
ART	Antiretroviral Therapy
BMI	Body Mass Index
CD4	Cluster of Differentiation 4
CDC	Centres for Disease Control
CI	Confidence Interval
HAART	Highly Active Antiretroviral Therapy
HIV	Human immunodeficiency virus
HR	Hazard rate
INLA	Integrated nested Laplace approximation
IQR	Interquartile Range
MSE	Mean squared error
PHRU	Perinatal HIV Research Unit
PMTCT	Prevention of Mother-To-Child Transmission
TB	Tuberculosis
UNAIDS	Joint United Nations Programme on HIV and AIDS
WHO	World Health Organisation

Preface

HIV remains a burden in sub-Saharan Africa where most of the infected people live. While the risk of mortality is lower due to the availability of free antiretroviral therapy in most sub-Saharan countries, a significant number of the infected people continue to die. Furthermore, those dying are largely in the productive stages of their lives. Hence further research is required to provide further insights on the factors associated with mortality in this cohort especially in patients on long term treatment.

Working in the field of HIV and TB research, I have had an opportunity of analysing several datasets for different purposes including conference abstracts, regulatory authorities and manuscripts. In the process, I was confronted by interesting situations where statistical modelling alone was insufficient in explaining study findings. While statistical analysis is helpful in assessing a hypothesis, there may arise situations where “the why” part of the findings cannot be sufficiently explained by modelling.

We sought to analyse mortality risk factors in HIV-infected individuals using Cox regression analysis modelling with frailty. Frailty is simply a term that assesses heterogeneity within a cluster or at the individual level in a sample. In our study, clusters were defined as either clinics where patients sought care or individuals with longitudinal data collected during follow-up. Hence in this study, frailty was assessed at the cluster level for the simulations and repeated measures data for individuals enrolled in the Adult Wellness cohort study. Both are indications of the shared frailty model assuming similar frailty within a cluster (or set of observations in repeated measures in an individual). Within an individual, the set of repeated measures form a cluster. Under the guidance of my supervisors, we sought to critically review literature on frailty modelling specifically applied to HIV-infected cohorts during the

period 2002-2011. In the literature, a low number of publications had utilised frailty modelling in assessing HIV risk factors. Furthermore we found no single publication assessing HIV risk factors using a mixed methods approach that incorporates frailty modelling and qualitative research. Following these findings and discussions with my supervisors, a clearer picture of a potential research question for PhD emerged. This was further strengthened by establishing key objectives that generated at least three publications.

The benefits of using methods that complement each other (mixed methods) are illustrated when assessing risk factors for mortality in HIV-infected people.

> *"Mixed methods research is important in health systems because it allows researchers to view problems from multiple perspectives, contextualize information, develop a more complete understanding of a problem, triangulate results, quantify hard-to-measure constructs, provide illustrations of context for trends, examine processes/experiences along with outcomes and capture a macro picture of a system" (Creswell and Clark, 2011)*

1.0 Introduction

1.1 Problem statement

In an ideal study, researchers would just collect as many relevant variables as possible that may help in addressing a particular research question. However, due to funding and logistical constraints, this is often impossible. Depending on the research question at hand, this may pose challenges during the analysis - requiring the use of a combination of methods in addressing the primary objectives.

Research studies generally involve the collection of data from people with similar characteristics such as those residing in the same household or neighbourhood or attending the same school or health clinics.[1-5] In other instances, research studies may involve collection of data longitudinally on the same people over time. In these cases, data collected within the same setting or from the same individuals during follow-up are correlated and are assumed to form a cluster. Other examples of clustering include healthcare workers employed in the same facility as well as participants enrolled in the same centre in a multi-centre clinical trial. Evaluating the health outcomes of people clustered in the same setting may involve analysing a time-to-event outcome like death or birth in a family or the first incident of a specific disease or hospitalisation. Such outcomes may be clustered within a household, a community, a school or a hospital due to similarity between members leading to correlation in cluster units.[1, 2, 4-8] Additionally, there may be other important variables that were not observed or collected during the study as well as variations between members of a cluster. The variations may involve some members being sicker than others hence introducing heterogeneity within the cluster.

Currently, the analysis of time-to-event data with mortality as outcome in HIV-infected individuals' involves using classical statistical approaches like Cox

proportional hazards regression.[9-12] However, the assumption underlying this methodology such as existence of a uniform risk[13] in mortality, ignores the effect of unobserved variables as well as participant heterogeneity within clusters or across participants in the same cluster. Ignoring these effects may lead to biased estimates and conclusions. Hence modelling mortality in data such as this requires application of suitable techniques that will account for the effect of unobserved heterogeneity.

Frailty models may be individual (unshared) or shared. In the individual frailty models sometimes referred to as univariate frailty, diversity or heterogeneity of data is evaluated within individuals; in the shared frailty model, sometimes referred to as multivariate frailty, groups of individuals such as patients in a clinic are assumed to share the same frailty.

More complex survival analysis techniques that form extensions of the classical Cox regression analysis exist that may be suitable for fitting time-to-event models that account for unobserved variables as well as within cluster heterogeneity. This includes Cox proportional hazards regression modelling with frailty[1, 4, 5, 7] that adjusts for unobserved heterogeneity. The approach has the capability of quantifying the effect of unobserved heterogeneity while at the same time indicating whether the risk of an event increases or decreases significantly in the presence of unobserved variables. Survival frailty models are often applied using different statistical distributions for the frailty term. Several possibilities for the frailty distribution exist and include making assumptions about the type of distribution (parametric) or working under a distribution free (semiparametric) framework. The performance of the frailty methods is tested through simulated data that compares them. Statistical distributions are used to generate data that mimic's the real world which can then be used for analysis. Simulations are powerful and important since they are less biased

and lead to plausible parameter estimates during the analysis. Through simulations, this study sought to identify the most suitable technique for fitting the Cox regression analysis with frailty and accounting for unobserved heterogeneity in a cohort of HIV-infected adults.

When unobserved heterogeneity is found to exist, the variable is often unknown because it was never collected during the study. However during the analysis where frailty is assessed, a cluster variable such as a clinic or individual with longitudinal data is specified. In this study, clusters representing clinics were assessed for frailty in the simulations whereas in the HIV Adult Wellness cohort data, individuals with longitudinal data were used as the clustering unit. Subsequently, it is explained by using speculative but plausible variables. In cohort studies involving HIV-infected individuals, speculative reasons may not truly reflect the true situation on the ground. This therefore justifies the need for a qualitative follow-up study to provide further insights that complement the quantitative analysis findings using the mixed methods approach. The qualitative findings are used to explain the unobserved heterogeneity (represented by the frailty result) from the quantitative analysis. There are few studies that use a mixed methods approach for investigating risk factors for mortality in HIV-infected people. Hence, this study proposes to explore the benefits of integrating both quantitative and qualitative analysis in identifying mortality risk factors. It is envisaged that findings from the mixed methods approach could provide further details that could be utilised to improve patient care and management.

1.2 Study justification

South Africa is the most successful country in providing free anti-retroviral treatment to HIV-infected people in the public sector. It is also among the countries with the highest burden of HIV and tuberculosis as well a good coverage of treatment for both diseases.[14] The free ART programs have been scaled up by the government in order to initiate therapy to as many eligible people as possible following the treatment guidelines. Once ART treatment is initiated, an HIV-infected person remains on treatment for life thus requiring appropriate patient care and management. In the long term however, there is need to understand these risks better for preventative purposes and identification of modifiable factors. As such, several studies in literature report on risk factors for mortality in HIV-infected people using short-term studies.[9, 10] Longer-term studies where conditions might change over time are likely to identify modifiable variables of interest that may be important in the care of patients. Such can be identified through use of robust methods that include both quantitative and qualitative techniques, otherwise called mixed methods, if utilised complementarily. In quantitative analysis involving time-to-event data, survival frailty modelling is a statistical modelling approach that enables one to account for unobserved variables. Furthermore, it also quantifies the magnitude of the unobserved measure. This improves precision of estimates for the predictors of mortality through alleviation of bias.[6, 15, 16]

While quantitative methods are popular in assessing risk factors for mortality, they lack a strengthened justification of their findings by not using a mixed methods approach. The use of mixed methods analysis provides the benefit of combining findings from quantitative and qualitative approaches. These findings are triangulated to present an integrated set of findings that are more informative than

those deduced from one approach. Mixed methods approaches are slowly gaining ground in the research community due to their power in providing complementary results.

1.3 Aims and objectives of the study

1.3.1 Overall aim

To determine mortality risk factors in HIV-infected people through incorporating unobserved variables using a mixed methods approach in which quantitative findings are explained by the qualitative.

1.3.2 Specific Objectives

1. To critically review statistical methods used for assessing risk factors for mortality in HIV-infected people between the years 2002 and 2011.
2. To use parametric and semiparametric frailty model simulations and real HIV data to compare mortality risk factors using a variety of methods and identify the type of model that provides a better fit.
3. To assess mortality risk factors in HIV-infected people in South Africa in different settings using the model identified in objective 2.
4. To use mixed methods (combining the best fitting model applied on retrospective data and qualitative analysis on prospective data) to evaluate mortality risk factors.

1.4 Background

The burden of HIV/AIDS mortality remains a major challenge worldwide despite the introduction of free anti-retroviral treatment (ART). According to the UNAIDS 2010 AIDS Global report, the number of HIV/AIDS related deaths has reduced from approximately 2.1 million in 2004 to 1.8 million in 2009.[17] This is largely attributed to the availability of free ART, care and support and decreasing incidence rates as a result of successful prevention of mother to child transmission (PMTCT).[17] Other intervention measures like behaviour change, condom promotion and distribution and medical male circumcision have also played a role in reducing HIV infection rates.

Globally, two out of three HIV-infected people reside in Sub-Saharan Africa, the region with the highest burden of the disease. The prevalence of the disease varies from under 1% to above 20%. But some of the countries with low prevalence have large populations translating to large numbers of people who are HIV-infected. Just above 11 million HIV-infected people were residing in Southern Africa in the year 2009, a figure that is higher than that reported a decade earlier.[17] Though HIV incidence rates have dropped in Sub-Saharan Africa, the proportion of new cases remains high.[18] It is estimated that approximately 1.2 million people died in 2010 whereas the number of new infections was significant.[17, 18] About half of the global mortality figures associated with HIV in the year 2010 were traced to Southern Africa.[18]

At 5.6 million estimated HIV-infected people, South Africa is among the countries with the highest proportion of HIV positive people globally even though they have successfully implemented PMTCT programs with a coverage of above 90%.[17, 18] Though ART has reduced mortality rates in general, a significant number of deaths

are attributed to HIV.[18-20] Investments by the government and donors have resulted in significant changes in the management of HIV in the country leading to an increase in prevention programs and widespread availability of ART that have lowered mortality associated with HIV.

With the scaling up of ART programs in many Sub-Saharan African countries, further insight on the long-term implications of treatment and its effect on mortality is required in order to provide optimal care. A plethora of literature exists on risk factors for mortality in HIV-infected people.[10, 21-23] Such literature has however seldom utilized robust statistical methods that provide a better assessment of mortality risk factors and the effect of hidden (latent) variables.

Statistical analyses (quantitative technique) of such data often use classical approaches such as logistic and Cox regression (survival analysis) models. Logistic regression models require large sample sizes to ensure sufficient power to detect mortality differences especially in checking model fit.[24-26] Cox regression makes the assumption of uniform risk in mortality across all patients (homogeneity).[13, 27] This assumption is generally limited because it is known that frail patients have a higher chance of death.

This study seeks to investigate risk factors for all-cause mortality in HIV positive participants using a mixed methods approach. The quantitative analysis proposed will be an extension of the classical survival analysis. The study will compare several survival analysis methods including frailty models using simulations and real data. We postulate that if properly modelled, frailty models provide a better fit compared with classical Cox proportional hazards regression. Thereafter, a qualitative analysis will be conducted involving participants with a deceased HIV-infected relative or friend. Findings from the qualitative and quantitative analysis will be used

complementarily to better understand the risk factors for mortality in HIV-infected people. Qualitative results will specifically be used to provide further insight to the frailty modelling results.

1.5 Conceptual framework

In order to show the process that was followed, Bhaskar's critical realism conceptual framework was applied.[28] This framework advocates for the use of a combination of methods in the analysis of research data with the understanding that they complement each other.[29-31] The study was designed to use quantitative analysis in assessing the literature review of publications on risk factors for mortality in HIV-infected people in a specified period followed by analysis on simulations and real data. The literature review involved identifying the commonly used statistical techniques in publications reporting on mortality risk factors. Thereafter simulations were used to identify the best fitting model that was later applied to the analysis of the Adult Wellness data. This was later followed by a qualitative analysis whose findings were used to explain those from the quantitative. As a whole, combinations of methodologies were used including in-depth analysis of qualitative data in explaining the findings as advocated by the critical realism framework.

We assessed the pathways presented in Figure 1 for the objectives specified above and the results are presented as attached publications at the end of this thesis. The literature review publication (see attached original paper 1) identified the methodological gap while simulations were helpful in identifying the best fitting and most robust method (original paper 2). The evaluation of mortality risk factors by setting (rural versus urban) identified the presence and quantified the effect of the unobserved (latent) variables (original paper 3). Findings from the qualitative analysis were instrumental in complementing those from the quantitative analysis by

providing further insight on the nature of important unobserved variables that may be affecting mortality which were not captured in the Adult Wellness study (paper 4 in progress).

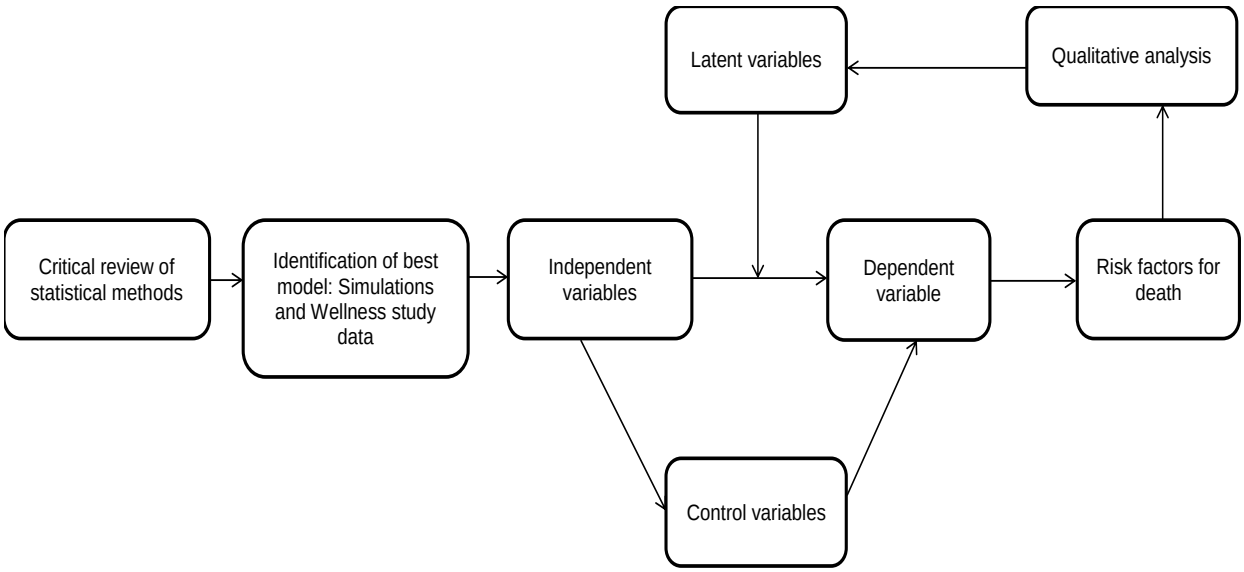


Figure 1: Critical Realism Conceptual Framework

2.0 Literature review

2.1 A critical review of existing methods

In biomedical research, the use of appropriate statistical methods in the analysis of study data is increasingly receiving more attention. This is demonstrated by the inclusion of statisticians in the review of manuscripts. A plethora of publications exist on the design of studies and suitability of the statistical analysis techniques applied.[32-36] Incorrect methodology may lead to wrong results and conclusions that may lead to poor decisions in the management and care of patients.

Reviews on commonly used statistical methodologies suggest use of simple inferential analysis methods that are easy to implement in standard software. A common methodology for assessing risk factors for mortality is the logistic regression model,[22, 37, 38] where predictors of a binary response are determined.[26] While the implementation of logistic regression is fairly straightforward, it ignores the effect of time especially in long-term studies with heterogeneity between patients.

Cox proportional hazards (Cox-PH) regression is widely used in evaluating time-to-death because it has the benefit of incorporating the effect of time and assumes that the risk of death is uniform across participants.[13] Research assessing risk factors for mortality favour this method.[9-12, 39] However, participants with a longer term of follow-up experience varying levels of disease and mortality risk (heterogeneity). Unfortunately, the Cox-PH model ignores heterogeneity and the relatedness between measurements taken from the same individual.

While many publications exist on risk factors for mortality in HIV-infected individuals, reviews assessing design of studies, nature of statistical analysis techniques used and their suitability is sparse. As such, there is need to understand the robustness of the methods used for the benefit of the public health community. Therefore, we

evaluated original publications meeting our study criteria in HIV-infected cohorts in the period 2002-2011 to assess the suitability of the methods used.

2.2 Comparison of survival analysis methods

While Cox-PH is the most implemented survival analysis technique, several extensions have been developed. In a review of oncology studies, the majority provided inadequate summary of methods.[40] The most commonly used methods were the logrank test and the Cox-PH. In a review of the survival analysis methods for assessing competing risks (defined as mutually exclusive events), variants of the Cox-PH model were implemented. The effect of incorrect assumptions on parameter estimates and confidence intervals have previously been evaluated.[41] When evaluating competing risks events in survival analysis of inferiority studies, the use of multiple time-to-event methods provides an opportunity to assess robustness.[42] Methods such as competing risks, marginal, frailty and Cox-PH models are recommended. Non-parametric covariate adjusted survival models perform better in non-inferiority studies compared to Cox-PH models.[43]

When the assumption of uniform risk of mortality (homogeneity) is violated, the Cox-PH model provides a poor fit. In this case, modelling focuses on addressing heterogeneity using the Cox-PH with frailty. Frailty models are superior when heterogeneity is inherent in data and perform better than the classical Cox-PH.[4, 5, 7, 42, 44, 45] Though more complex to fit, the Bayesian frailty models perform better.[1, 7, 45, 46]

Frailty modelling assumes some form of distribution for the frailties. Several distributions for the frailties have been proposed in the literature with the most common being the gamma and log-normal. The gamma distribution is commonly used due to its mathematical tractability while the log-normal fits in well with the

generalised linear mixed models. Frailty models may be implemented in parametric or semi-parametric forms.[47] In the parametric form, a distribution is assumed for the survival time while in the non-parametric form, no distribution is assumed allowing for flexibility in the modelling process. While methods for modelling the parametric form are well developed, the semi-parametric form requires use of more complicated approaches. [47, 48]

Semi-parametric frailty models have been used to compare various frailty distributions with the most common frailty distributions being the gamma and log-normal. Using a comparison of four models involving correlated data, it was shown that the semi-parametric frailty models perform better than the naïve Cox model. [49] The use of the Dirichlet process prior in the Bayesian modelling of frailties has been shown to perform better in the semi-parametric frailty modelling compared with the parametric. [50] Using simulations, several software packages capable of fitting shared semi-parametric frailty models were compared and found to fit well but with varying levels of complexity. [2] Other examples on the performance of various semiparametric methods have been presented in literature using both gamma and lognormal frailties. [51-53]

Though frailty models are not commonly used in epidemiological research evaluating risk factors for mortality in HIV-infected people,[54] they provide better, more reliable and robust estimates. We simulated survival data and analysed it using five methods available in standard software: Cox-PH, time varying Cox-PH, shared Gamma frailty with Weibull and exponential baseline hazards and the integrated nested Laplace approximation (INLA)[17, 46, 55, 56] with a Weibull baseline. INLA is a recent Bayesian inference technique that provides accurate approximations and is computationally fast.[45, 46, 55] Simulations were implemented to enable the identification of the best

fitting model through evaluating the fit statistics, parameter estimates and standard errors. The mean and standard deviations were similarly assessed for INLA. To our knowledge, studies comparing frequentist and the Bayesian INLA time-to-event methods are limited. Furthermore, our findings are an addition to the existing literature comparing frequentist and Bayesian time-to-event frailty models. In addition, articles presenting an overview of maximisation methods in semiparametric frailty models that include the integrated nested Laplace approximation are limited. This study provides such an overview using the gamma and lognormal frailties through simulations.

2.3 Risk factors for mortality

Mortality rates in HIV-infected people in Africa are declining because of free ART in public healthcare facilities. However, despite the availability of free treatment, patients continue to start treatment when they have advanced HIV disease, with a significant number succumbing to death soon after initiating treatment. As a result, it is imperative to re-assess risk factors for mortality with the aim of identifying where the challenge lies.

Commonly reported risk factors vary across the literature and include clinical and socio-economic factors. Clinical factors include: declining CD4 count and increasing viral loads;[10, 15, 23, 57] World Health Organisation (WHO) stages;[23, 39] Centres for Disease Control (CDC) clinical staging;[10] low haemoglobin levels;[21, 23] and opportunistic infections[6, 11, 58] such as pneumonia and undiagnosed and untreated tuberculosis (TB). Demographic variables such as age and gender have been associated with mortality.[6] Education level, smoking of tobacco products[59] and use of alcohol has similarly been associated with an elevated mortality risk in persons with HIV.[58]

The care of HIV-infected patients is improved by an in-depth understanding of the context in which mortality is likely to occur and how it can be mitigated. Different factors in both rural and urban settings in developing countries affect access to care impacting on the timing of treatment initiation or care seeking. Furthermore, the long term implications of anti-retroviral treatment need to be well understood. To understand these dynamics, epidemiological studies are required in African countries that bear the greatest burden of the disease and other resource limited countries where coverage of free antiretroviral care programmes has been rolled out. Analysing epidemiological research data using optimal statistical methods that incorporate unobserved variables may be helpful in providing further insights in the management and care of HIV patients. Additionally, these methods should seek to identify modifiable factors that may be used to develop tailored interventions.

3.0 Methodology

3.1 Study setting and population

In the quantitative analysis, data from the Adult Wellness study, a cohort of HIV-infected adults, was analysed.[58] Participants were enrolled at the Soweto based Perinatal HIV Research Unit (PHRU) and Tintswalo hospital in Mpumalanga, South Africa. The Adult Wellness study, whose data was used for the quantitative analysis in this study, was conducted from the year 2003 and ended in the year 2010. Thereafter, a qualitative study was conducted at PHRU between May and August 2014 targeting participants who had reported a bereavement involving a person who was HIV-infected.

PHRU is a leading clinical research centre located in a more urbanised area while Tintswalo is a primary healthcare facility located in a rural area. Both sites managed patients using the same treatment guidelines as provided by the South African government.

The township of Soweto is in the Southwestern part of Johannesburg and is estimated to have a population above a million people.[60] It is estimated that a large proportion of the working population in the city of Johannesburg reside in Soweto. Poverty levels and high unemployment rates are the biggest challenges in the area.[60]

The overall HIV prevalence of the province where Soweto is located is 12.4% and females are disproportionately affected by HIV.[61] Access to free anti-retroviral treatment has reduced mortality rates associated with HIV but challenges such as delays in seeking care and starting medication timeously as well as development of opportunistic infections such as TB remains.[62-64]

Tintswalo is a rural based hospital that is located in Mpumalanga province of South Africa. It serves the people of Bushbuckridge district whose population size is about 500, 000 people. High poverty levels affecting approximately three-quarters of the population have been documented in Bushbuckridge district. [65]

3.2 Ethical consideration

The Human Research Ethics Committee (Medical) of the University of the Witwatersrand provided approval to conduct the study (clearance number M130818) as well as for the Adult Wellness study (clearance number M021129). All the participants willingly provided written informed consent to participate in the study.

In the qualitative study, consenting was provided for study participation and audio-recording. The consent forms and interview guides were written in English but the consenting process and the in-depth interviews were carried out in either English or Zulu as appropriate. Participants were re-imbursement (about $15 that approximated to 150 Rands at the time) for their time. To reduce selection bias, reimbursement was only made known during consenting. Counsellors were available at the PHRU ZAZI VCT clinic to provide counselling to any distressed participants.

3.3 Data and methods for objective 1

Objective one involved a critical review of existing literature drawn from papers written on mortality risk factors in people with HIV. The areas of focus were in the design of the study, the methodology used for analysis and their suitability.

3.3.1 Search strategy and selection criteria

The online reference library Pubmed/Medline was searched for the terms ‘‘Predictors of HIV Mortality’’, ‘‘Determinants of HIV Mortality’’ and ‘‘Factors associated with HIV mortality’’.[54] Publications had to be written in the English language to be eligible. Only publications written in the period 2002-2011 were retrieved as this period coincided with the window the Adult Wellness study was conducted. To evaluate whether an improvement in the use of statistical methods occurred over time, outcomes were analysed in the periods 2002-2006 and 2007-2011. Our criteria focussed primarily on original articles while excluding others such as editorials.

Data were collected following a modified version of the method of Colditz and Emerson[66, 67] and put on to a spreadsheet. The statistical methods extracted from the eligible publications were classified variously as outlined elsewhere.[54]

3.3.2 Data analysis

The data were categorised into two periods 2002-2006 and 2007-2011 and comparisons made between them. Since all the data collected was categorical, frequencies were determined followed by chi-square analysis or Fishers exact test as appropriate. All the statistical analysis was assumed to be two-sided and performed using SAS Enterprise Guide 5.1 (SAS Institute, Cary, NC) at 5% level of significance.

3.4 Data and methods for objective 2

In the second objective, the parametric and non-parametric frailty models were used to analyse simulated data. Parametric methods are presented first followed by the non-parametric.

In the parametric frailty modelling, four survival analysis techniques were used: Cox proportional hazards regression (Cox-PH), shared Gamma frailty (with Weibull and exponential baseline hazards) all frequentist and the Bayesian integrated nested Laplace approximation (INLA) with a Weibull baseline hazard.

These methods were also applied to real data from a HIV-infected cohort in South Africa.

3.4.1 Parametric frailty model simulations

The four methods were compared through a simulation. Data were simulated from the shared frailty model $h_{ij}(t) = h_0(t)w_i \exp(X_{ij}\beta)$ where, $h_0(t)$, w_i, X_{ij} and β where $h_0(t)$ is the baseline hazard, w_i is the frailty term, X_{ij} are the observed covariates for subject j in cluster i and β are the parameter estimates. The Weibull distribution, defined as $h(t) = \lambda t^{\alpha}$ with a uniform exponential baseline hazard was utilised for modelling the hazard for event and censor times. We first generated the frailties,

defined as $w_i \sim N(0, \vartheta)$ where ϑ is unknown variance, that were determined across clusters and number of participants per cluster. [68] The observed time was considered as the minimum of (T, C) where T and C are the event and censoring times respectively. *T* and *C* were derived from a Weibull distributed linear predictor to generate event and censoring times. The covariates X_{ij} (gender, employment status and CD4 categories) were generated from the uniform distribution $U \sim (a, b)$ after setting some parameters derived from the Adult Wellness study using the Cox proportional hazards regression. Categorisation (e.g male and female) of the variables in the simulated data was done using proportions determined in an actual HIV dataset on completion of follow-up. Participant information was assessed at their last visit. We assessed the probability coverage, the interval containing the true value most of the time, of the simulated data at 95% confidence level.

We ran 1,000 simulations of a hypothetical data with 5, 10 and 20 clusters each of cluster size 50 and 100 participants. The final estimates/mean and standard errors/standard deviation were averaged over the simulated data.

3.4.2 Statistical models

Notation

Let $j = 1, \dots, n$ participants be enrolled in a study with $i = 1, \dots, n_j$ clusters and a time-to-event endpoint. We assume that participants are followed until time T_{ij} when an event occurs or consent is withdrawn or are lost to follow-up or study completion. Let the observation time for participant j in cluster i be defined as $T_{ij} = \min(T, C)$ where $\sigma = 1$ is the censoring variable defined as:

$$\sigma = \begin{cases} 1 : if\ T_{ij} \leq \sigma_{ij} & The\ observation\ time\ is\ not\ censored \\ 0 : if\ T_{ij} > \sigma_{ij} & The\ observation\ time\ is\ censored \end{cases}$$

Let X_{ij} be the explanatory variables such as gender, employment status and CD4 count. Hence we may now define the hazard for person j in cluster i as $h_{ij}(t)$. For the purpose of illustration and without any loss of generality, analysis is confined to the covariates X_{ij} outlined above.

We compare four selected time-to-event methods available in standard software: Cox-PH, shared Gamma frailty with Weibull and exponential baseline hazards and the integrated nested Laplace approximation (INLA) with a Weibull baseline hazard. The study site is considered a clustering variable in which frailty modelling is used to assess unobserved heterogeneity.

3.4.2.1 Cox proportional hazards (PH) regression

The Cox-PH is a semi-parametric regression model defined as $h_j(t) = h_0(t)e^{X_j\beta}$ where $h_0(t)$ is an unspecified baseline hazard, $\boldsymbol{X}$ is the vector of covariates and β are the parameter estimates.[13, 69] The parameter estimates $\beta_{j's}$ are determined through the likelihood function $L(\theta|y) = \prod_{j=1}^{n} p(y_j|\theta) = \prod_{j=1}^{n}[f(t_j)]^{c_j}[S(t_j)]^{1-c_j}$ on adjusting for censoring.[69] $f(t_j)$ is the probability density function while $S(t_j)$ is the survival function at a specified time and C_j is the censoring indicator for subject j. The hazard $h_j(t)$ in this case allows for a population-averaged interpretation of the effect over all the clusters.

3.4.2.2 Weibull Parametric Cox regression with Shared Gamma frailty

To account for the effect of unobserved variables, the hazard function is modified to: $h_{ij}(t;x) = h_0(t)e^{X_{ij}\beta+Z\gamma} = h_0(t)w_i e^{X_{ij}\beta}$ where w_i is the Gamma frailty term (with a known mean and unspecified variance) and $h_0(t)$ is the Weibull-distributed baseline hazard $h(t) = \lambda v t^{v-1}$. w_i accounts for unobserved heterogeneity in the risk of death through estimating the variance. [5, 7] Large values of the variance σ^2 reflect a greater

degree of heterogeneity. The Gamma frailty model is the most popular due to its simplicity in computation and analytical form in survival and hazard functions. [1, 4, 5, 7, 44]

3.4.2.3 Exponential Parametric Cox regression with Gamma frailty

In the frailty model $h_{ij}(t;x) = h_0(t)w_i e^{X_{ij}\beta}$, the exponentially-distributed baseline hazard $h_0(t) = \lambda$ has a constant hazard.

3.4.2.4 Bayesian Frailty modelling using Integrated Nested Laplace Approximation (INLA)

Following the development of INLA using Bayesian inference, [45, 46, 55, 56] the hazard rate for the *i*-th subject is defined as $h_i(t;x) = h_0(t)\exp(\eta_i)$ where n_i is a structured additive predictor from Latent Gaussian models. Using a baseline hazard function derived from the Weibull distribution, $h_i(t;x) = \lambda v t^{v-1}\exp(\eta_i)$ and $\eta_i = \beta_0 + \beta_i + \log w_i$ where $\beta_0, \beta_i \sim N(0,0.001)$ and $v \sim \Gamma(1,0.001)$ are prior densities. w_i are frailty effects $\sim N(0,\tau^{-1})$. The likelihood is assumed exponential and defined as $\prod(y_i|\eta_i,\theta)$ where θ is the set of hyper-parameters. Parameter values are determined from estimates of the posterior marginals $\hat{p}(x|y)$ and $\hat{p}(\theta|y)$ that are resolved numerically in INLA. [45]

3.4.3 Parametric Frailty Model Data Analysis

For the Cox-PH and frailty models, data were assessed at the end of follow-up. In the frailty models, site was considered as the clustering variable. The Gamma frailty was modelled with a Weibull and exponential baseline hazard while INLA, the Bayesian frailty, was modelled with a Weibull baseline hazard.

For comparison across the frequentist models, -2 loglikelihood fit statistics were used. Parameters and standard errors (frequentist models) as well as means and

standard deviations (from INLA) were compared to assess their similarity. INLA generates the Deviance Information Criterion (DIC) defined as $DIC = \bar{D} + pD$ where $\bar{D} = -2 * loglikelihood$ [70, 71] and pD is the number of effective parameters. The mean squared error (MSE) defined as $MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$ was used to compare parameter estimates across all the simulated models excluding the model with five clusters.

Cox-PH was fitted using SAS Enterprise Guide 5.1 (SAS Institute, Cary, NC) while Gamma frailty models with Weibull and exponential baseline hazards were determined in STATA 14. [72] INLA was assessed by the R statistical software. [73]

3.4.4 Non-Parametric frailty model simulations

Simulations for the semiparametric frailty modeling were conducted using the frailtysurv R package approach. [74] Underlying assumptions were the mean and variance of the gamma distribution was $E(X) = 1$ and $Var(X) = \theta$ whereas the lognormal was $E(\theta/2) = 1$ and $Var(X) = E(2\theta) - E(\theta)$. For each of the frailty models, 10 clusters of size 50 and 100 were generated 1,000 times. For each j^{th} observation that belongs to the i^{th} cluster, the conditional survival function was derived by $S_{ij}(t|Z_{ij}, w_i) = exp(-H_0(t)w_i exp(\beta X_{ij}))$ where the unspecified cumulative hazard function is represented by $H_0(t)$ whereas w_i and βX_{ij} are as previously defined [74]. To simulate the failure time T_{ij} in the region $(0, \infty)$, $H_0(t)$ is inverted and the failure times are determined by $T_{ij} = H_0^{-1}(-ln(U_{ij})\exp(-\beta X_{ij})/w_i)$ where U_{ij} is uniformly distributed. The generated frailties are dependent on the selected frailty distributions, in our case the gamma and lognormal. Using a normal distribution for right censoring, an indicator for failure time was defined as $\sigma_{ij} = I(T_{ij} \leq C_{ij})$ where

C_{ij} is the censoring variable for the j^{th} subject in the i^{th} cluster. In our simulation, the censoring rate was fixed at 20%.

3.4.4.1 Penalized partial likelihood

In this approach, smoothing splines that help in parameter estimation through transformations are used in conjunction with the penalized partial likelihood[75, 76] to determine parameter estimates. The penalized partial likelihood is defined by $l^{s}_{PPL}(\beta) = log\left[\prod_{i=1}^{n}\prod_{j=1}^{n_i}\left[\exp\big(g(.)\big)/\sum_{R(t_{ij})}\exp(g(..))\right]^{\sigma_{ij}}\right] - \lambda\int\{g''(z)\}^2dz$ where the parameter estimate β defines functions $g(.)$, $g(..)$ and $g''(z)$ and $\lambda\int\{g''(z)\}^2dz$ is the penalty term subtracted from each iteration during the parametrization process.[75, 76]

3.4.4.2 H-Likelihood

The h-likelihood functions fit both the gamma and lognormal frailty distributions. The approach utilises the Laplace approximation procedures leading to efficient parametrisation of frailty estimates. The approach is defined as $h = h(\beta, h_0, \alpha) = l_0 + l_1$ where l_0 and l_1 is conditional log densities and sum of log densities respectively. [77] An in-depth analysis of the h-likelihood theory and its potential applications are presented elsewhere. [77-79]

3.4.4.3 Piecewise log-constant baseline hazard

In this Bayesian approach, the time of observation is partitioned and the baseline hazard in each partition remains constant. The covariate effect is assumed uniform as well. Here, parameter estimation is carried out using data augmentation through latent Gaussian models and through a first-order random walk (RW1). [80, 81] The likelihood of the j^{th} observation in the i^{th} cluster is given by $\pi\big(y_{ij}|\eta_{ij},\theta\big)$ where η_{ij} is a linear predictor and θ is a vector of hyper-parameters. [81]

3.4.4.4 Penalized full Likelihood

This method uses a full likelihood in conjunction with splines in parameter estimation. Splines are useful where precise estimations in semiparametric frailty modelling are unachievable [82]. A penalty is imposed on the likelihood by this method through introduction of large values on functions classified as rough. It is defined as $pl(h_0(.),\beta,\theta) = l(h_0(.),\beta,\theta) - \sum_{d=1}^{K} K_d \int_0^{\infty} h''^2_{0d}(t)\, dt$ and $l(h_0(.),\beta,\theta)$ is the full likelihood and $K_d \geq 0$ are smoothing parameters per stratum [82, 83].

3.4.5 Semiparametric Frailty Model Data Analysis

Simulated data were analysed by fitting semiparametric gamma and lognormal frailty models where parameter estimates, their standard errors and frailty variances were generated. For comparison, box plots were generated for parameter estimates and standard errors for the models. The mean squared error (MSE) was used to compare parameter estimates across all simulated models.

The semiparametric frailty models were fitted using the following R packages: Integrated Nested Laplace Approximation (*INLA*) for Piecewise log-constant baseline hazard, *frailtyHL* for H-likelihood, *coxPH* for penalized partial likelihood and *frailtyPack* for penalized full likelihood. The SAS SGPLOT procedure in SAS Enterprise Guide 7.1 (SAS Institute, Cary, NC) was used to generate plots.

Findings from the parametric analysis are not provided in detail as they are not considered original. Additionally, previous reviewers' comments indicated that non-parametric frailty modelling would be more original and worthy presenting. Hence manuscript 2 focusses on simulation results from non-parametric frailty modelling whose presentations are fewer in the literature of frailty modelling.

3.5 Data and methods for objective 3

The third objective involved modelling risk factors for mortality using the best fitting model identified in objective 2. Data were clustered at the individual level in the frailty models and risk factors for mortality were assessed by site.

3.5.1 Eligibility and Data collection

The Adult Wellness study was a nurse based program at both sites offering care to HIV-infected individuals. Funding was provided by PEPFAR through USAID South Africa and the program run from the year 2003-2010. Participants were referred into the program from surrounding clinics. Those eligible to start treatment under the South African guidelines were referred.

Being HIV positive, an adult ≥ 18 years and on care at PHRU or Tinstwalo hospitals as well as consent was required for enrolment. The study procedures are described in detail elsewhere.[84, 85]

Following enrolment, participants were followed up approximately every 6 months. Those missing visits were followed up through telephone calls and home visits. Data were collected on socio-demographic and clinical factors at enrolment. During follow-up, clinical measurements such as CD4 counts and changing socio-demographic measures such as employment were collected. Tuberculosis was confirmed if one of the following conditions was met: on multidrug therapy for TB, a positive outcome in the acid fast bacilli test, a positive mycobacterium culture test, hospital admission or death due to TB.

3.5.2 Measures

The primary outcome of the study was all-cause mortality that was confirmed by a home visit or contact with a relative. Other measures collected included socio-demographic variables such as age, gender and household income and clinical

measures such as body mass index (BMI) and the WHO stage. BMI was categorised following the CDC classification: underweight (< 18.5), normal (18.5-24.9) and overweight or obese (≥ 25). CD4 count was classified into the groups: 0-200, 201-350, 351-500 and > 500 cells/mm^3. The variables CD4 Count and BMI were treated as time-varying covariates. Further details on the methods and measures are presented in original paper 3 that appears as an attachment at the end of this thesis.

3.5.3 Statistical analysis

All the analysis was stratified by site: rural and urban. Descriptive statistics such as medians and interquartile ranges were determined for continuous measures such as age, CD4 count, BMI, crowding index and household income. These measures were compared between sites using the Kruskal-Wallis non-parametric test. Continuous measures such as age and BMI were categorised into age-groups and BMI groups respectively. Frequencies were determined for all categorical variables. Chi-square analysis and Fishers Exact test, whichever was appropriate, were used for the comparisons between sites.

Poisson regression modelling was used to compare mortality rates per 100 person-years between sites. Mortality rates were determined overall, pre-HAART and post-HAART. Graphical plots were generated for the comparison of rates by site. Kaplan-Meier graphs were plotted to show the hazard rate for mortality by site. Hazard rates were determined by Cox proportional hazards regression modelling.

There was missing information of approximately 18% and 11% in CD4 counts and BMI respectively. Missing data techniques were used to impute for incomplete data using the Multiple Imputation (MI) procedure in which five complete datasets were generated and analysed. [86-88] The average of the estimates was used as the final result.

The integrated nested Laplace approximation (INLA) with the shared gamma frailty model [45, 46] was used to assess mortality risk factors. Univariate analysis was performed first followed by multivariate analysis. Hazard rates and their 95% confidence intervals were determined through exponentiating model results. The multivariate analysis was conducted using variables that were thought important followed by application of the backward selection procedure. Variables thought to be important included age, CD4 count, BMI, crowding index, household income, length on HAART and an indicator for employment, smoking and whether they have ever had TB. Sensitivity analyses were conducted to assess different scenarios. This included assessing time to death overall and per site by CD4 count, BMI, gender, whether on HAART or not, length of time on HAART. Mortality rates pre and post HAART were also determined. Further details on the statistical methods are available in original paper 3 at the end of this thesis.

Frailty variance in both sites was determined through INLA. A frailty variance value > 1 suggests an increased hazard rate in the event of interest while that < 1 suggests a decreased hazard rate.

Statistical analyses were conducted in R (www.r-project.org) and SAS Enterprise Guide 5.1 (SAS Institute, Cary, NC).

3.6 Data and methods for objective 4

The fourth objective involves using a mixed methods approach in which findings from the qualitative analysis complement those from the quantitative as suggested by the critical realism framework. The methods for the qualitative research are presented in section 3.6.1.

3.6.1 Sampling

Participants were recruited through the PHRU ZAZI voluntary counselling and testing (VCT) centre. ZAZI is a successful HIV testing service that is offered for free to clients in Soweto and surrounding areas. In the present study, participants were purposively recruited from the ZAZI VCT. Of the 26 participants invited to participate, 20 were enrolled. During the VCT counselling process, clients were routinely asked about the loss of a person in their life and probed further to determine if it was due to HIV. Those responding in the affirmative were probed further by counsellors to determine who it was, when it happened and how they dealt with the loss. Thereafter, they were invited to enrol into the study. Interviews were scheduled within a period of seven days for those willing to participate. A day prior to the scheduled interview date, a telephone call was made to confirm the appointment.

3.6.2 Eligibility

To be eligible, participants had to be 18 years or older, have known a close person who had died due to HIV and were willing to participate in a face-to face in-depth interview.

3.6.3 In-depth interviews

Data were collected between May and August 2014 using in-depth interviews to understand the characteristics of the participant and the deceased persons. A demographic questionnaire assessed age, gender, primary home language and educational level and was administered prior to the in-depth interview.

Interviews were conducted in English and/or isiZulu by a trained interviewer who had experience in conducting qualitative interviews. The in-depth interviews explored participants perceptions about: the deceased persons' relationships with significant others or in general, health seeking behaviours, attitude of healthcare workers

towards HIV-infected patients seeking care, initiation and adherence to antiretroviral therapy (ART), home-based care support, nutrition and/or food security and financial stability. Interviews occurred in private rooms at the ZAZI VCT centre and lasted no more than an hour.

All interviews were audio-recorded and transcribed verbatim one day after the interview. Transcriptions were performed by an external team with experience in transcribing audio recordings that was approved by the study team. Where audio-recordings had a mixture of the local language and English, translations into English occurred during the transcription process. Every second transcript was selected by the interviewer for comparison with the audio recording to check for consistency. Where discrepancies were identified, audio-recordings were transcribed again. A total of 10 transcripts were selected for comparison with the audio-recordings for quality checks (that were done by a bilingual research assistant).

All the study procedures were recorded in a study work book that was used to monitor study progress. It contained information such as how many people had been interviewed, how many more were likely to be interviewed, scheduled interview appointments and administrative information such as participant re-imbursements for transport.

3.6.4 Data analyses

Data were analysed through a framework analysis approach where similar and interrelated codes were grouped together to form sub-themes and themes in a data matrix.[89, 90] Core codes were independently developed by two research assistants and agreed upon in collaboration with an experienced qualitative researcher. The first four transcripts were reviewed by two research assistants line by line and the findings used to develop a code book. A meeting was organised between the

research assistants and a qualitative analysis expert to verify the consistency of the codes. Divergent codes were discussed and a resolution reached on the most appropriate coding term. Subsequent transcripts were assessed based on the code-book. Updated codes were only included if subsequent codes could not fit within the initial ones.

The data matrix was instrumental in assessing the coding progress allowing for the development of sub-themes and themes.[89, 90] The process of developing sub-themes and themes was regularly refined during the analytic process as new codes emerged. Any discrepancies in coding were resolved through consensus. Development of codes, sub-themes and themes followed a combined approach of deductive and inductive data reduction procedures.

3.6.6 Trustworthiness of the data

Trustworthiness of the data was assessed using various methodologies.[91-94] An audit trail was kept to monitor the study progress such as evaluating the first two interviews and amending the interview guides as appropriate. Codes were independently prepared by the research assistants and harmonised in a meeting together with an experienced qualitative researcher. Thereafter, a codebook was developed for all subsequent interviews. Code development stopped when no further changes occurred. Audio-recordings were compared with the transcripts to verify consistency, ensure translations matched the recordings and make any appropriate amendments. The individual construction of the participants, which involves seeking clarification of their responses, was improved through probing by the research assistant. Final results were shared with the study team for verification of the themes and interpretation. Differences in the coding, themes and interpretations were resolved as a team.

4.0 Results

4.1 Review of methods for statistical analysis

Detailed findings on the design of studies, methodology used to analyse them and their suitability in the identified publications on mortality related risk factors in HIV positive people are presented elsewhere.[54]

4.1.1 Disposition flow chart

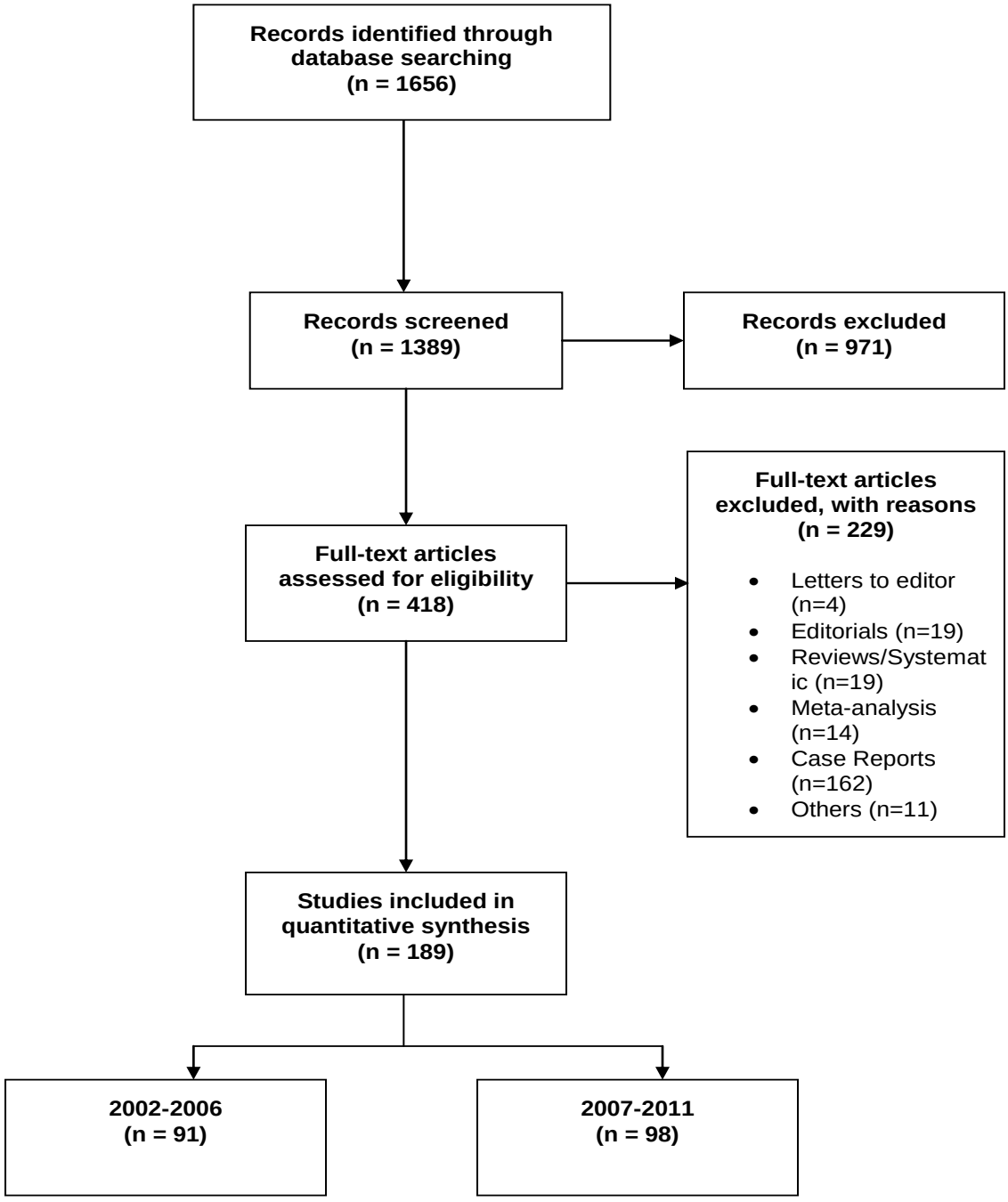


Figure 2: The flow of screened and selected publications

From 1656 initial publications, only 189 qualified for inclusion in this study. These were further split into two periods of assessment: the years 2002-2006 (n=91) and 2007-2011 (n=98). Articles were excluded if they were editorials, reviews (systematic and meta-analytic) and patient case reports. Figure 2 illustrates the process of selecting publications meeting the study criteria.

4.1.2 Design of studies

Table 1: Design of studies and study sample sizes

	2002-2006		2007-2011		
Research Design	**No.**	**(%)**	**No.**	**(%)**	**p-value**
Prospective	67	74	62	63	0.66
Retrospective	23	25	33	34	0.14
Case control	1	1	3	3	-
Prospective	1	-	0	-	-
Retrospective	0	-	3	-	-
Total	91		98		
Sample sizes					
up to 200	22	24	15	15	0.5
201 to 500	21	23	25	26	
501 to 1,000	16	18	20	20	
>1,000	32	35	38	39	

There were more prospective studies than retrospective studies and the proportions did not differ significantly between the two periods. A summary of the design of studies together with the number of enrolled participants is presented in Table 1.

4.1.3 Overview of statistical analysis approaches

A summary of the statistical methods commonly reported in the publications identified during the review are presented in Table 2. Of the 186 eligible publications, only 7 (4%) reported using frailty modelling. Though not significantly different between the periods, Cox proportional hazards regression was more commonly reported probably due to the study design.

Table 2: Summary of statistical methods

Statistical methods	2002-2006 No. of publications (N=91)	(%)	2007-2011 No. of publications (N=98)	(%)	p-value
Descriptive statistics	88	97	94	96	0.78
Inferential methods					
t-test	13	14	19	19	0.35
Contingency table analysis					
Basic (chi-square/Fisher exact)	25	27	37	38	0.13
Correlation	9	10	1	1	-
Epidemiologic statistics	39	43	45	46	0.67
Analysis of variance	3	3	1	1	-
Regression					
#Multiple regression	13	14	31	32	0.005
Survival analysis					
Kaplan-Meier	69	76	53	54	0.002
Cox-Regression	68	75	63	64	0.12
Time dependent Cox regression	21	23	11	11	0.03
Cox-Regression with frailty	1	1	6	6	-
Non-parametric methods	86	95	91	93	0.64
No. of different inferential methods					
2 or 3 methods	10	11	11	11	0.96
4 or 5 methods	44	48	43	44	0.54
More than 5 methods	37	41	44	45	0.56
¶Basic analysis	36	40	45	46	0.38
*Advanced analysis	90	99	97	99	0.96

#Includes logit and Poisson modelling
¶Includes parametric and non-parametric inferential analysis methods such as student t-test
*Includes logit, Poisson and time-to-event methods as well as epidemiological statistics

4.1.4 Suitability of utilised statistical techniques

There were 53 (78%) and 48 (77%) prospective studies from 2002-2006 and 2007-2011 respectively that reported using Cox proportional hazards regression (Table 3). Prospective studies reporting time dependent Cox proportional hazards regression were 16 (24%) in 2002-2006 and 10 (16%) in 2007-2011. Overall, frailty modelling (n=6) was more commonly reported in the period 2007-2011.

Table 3: Suitability of utilised techniques in assessing mortality risk factors

	2002-2006		2007-2011	
	Prospective	Retrospective	Prospective	Retrospective
	(N=68)	(N=23)	(N=62)	(N=36)
Statistical methods	**No. (%)**	**No. (%)**	**No. (%)**	**No. (%)**
Appropriate statistical analysis				
Cox regression	53 (78)	15 (65)	48 (77)	15 (42)
Time dependent Cox regression	16 (24)	5 (22)	10 (16)	1 (3)
Cox regression analysis with frailty	1 (1)	0 (0)	3 (5)	3 (8)
Poisson regression	3 (4)	0 (0)	4 (6)	0 (0)
Sub-Optimal statistical analysis				
Logistic regression	4 (6)	4 (17)	7 (11)	17 (47)
Unclear	1 (1)	2 (9)	2 (3)	3 (8)

NB: Some of the publications utilized ≥ 1 statistical method

4.2 Comparing time-to-event methods

4.2.1 Parametric frailty modelling

4.2.1.1 Simulation

The simulation results are presented in Tables 4, 5 and 6. To be consistent, all the four models are fitted using the same simulated dataset. In general, the simulation results across parameter estimates and standard errors as well as the INLA means and standard deviations are similar.

Within the frequentist models, the Gamma frailty models with Weibull (319.9 and 645.2 in table 4; 651.7 and 1300 in table 5; 1314 and 2609 in table 6) and exponential (314.9 and 651.4 in table 4; 652.8 and 1301 in table 5; 1315 and 2613 in table 6) baseline hazards had lower -2 loglikelihood values compared with the Cox-PH models (328.5 and 796.4 in table 4; 660.7 and 1600.5 in table 5; 1330.3 and 3215.4 in table 6).

The INLA model had a lower frailty variance (0.1 and 0.2 in table 4; 0.2 in table 5; 0.3 in table 6) relative to the Gamma frailty models with Weibull (0.3 in table 4 and 6; 0.4 in table 6) and exponential (0.3 in table 4 and 6; 0.4 in table 6) baseline hazards.

The probability coverage values were high for all the simulated datasets. In the data with five clusters, the coverage values ranged between 94.5% and 95.3% while it was 94.6% to 94.9% in the data with ten clusters. In the data with twenty clusters, the probability coverage ranged between 94.7% and 95%. The distribution of probability coverage for specific variables within different clusters is presented in table 7.

In the model with ten clusters (n=50 per cluster), the INLA MSE was lower for the parameter estimates for male gender and CD4 count between 201 and 350. In the model with twenty clusters, the INLA MSE was better in general (table 8).

Table 4: Five clusters of size 50 and 100

Variable	Cox Regression	Gamma Frailty (Weibull)	Gamma Frailty (Exponential)	INLA
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
5 Clusters (n=50 per cluster)				
Gender				
Male	0.8836 (0.3131)	0.8858 (0.3060)	0.9382 (0.3107)	0.8670 (0.3031)
Female	Ref	Ref	Ref	Ref
Employment Status				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.9861 (0.5071)	1.0177 (0.4694)	0.9437 (0.4891)	1.0680 (0.6050)
CD4 Categories				
0-200	Ref	Ref	Ref	Ref
201-350	-.3848 (0.3754)	-.3816 (0.3680)	-.3111 (0.3741)	-.3955 (0.3657)
> 350	-.8011 (0.4016)	-.8138 (0.3908)	-.7811 (0.4009)	-.8223 (0.4030)
-2 Log L	328.5	319.9	314.9	N/A
DIC	N/A	N/A	N/A	178.8
Frailty Variance	N/A	0.3	0.3	0.1
5 Clusters (n=100 per cluster)				
Gender				
Male	0.8782 (0.2143)	0.8800 (0.2116)	0.8691 (0.2105)	0.8756 (0.2117)
Female	Ref	Ref	Ref	Ref
Employment Status				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.9558 (0.3436)	0.9642 (0.3394)	0.9740 (0.3362)	0.9704 (0.3519)
CD4 Categories				
0-200	Ref	Ref	Ref	Ref
201-350	-.3961 (0.2586)	-.4008 (0.2564)	-.3851 (0.2540)	-.4006 (0.2561)
> 350	-.7851 (0.2730)	-.7938 (0.2710)	-.7783 (0.2688)	-.7947 (0.2708)
-2 Log L	796.4	645.2	651.4	N/A
DIC	N/A	N/A	N/A	267.9
Frailty Variance	N/A	0.3	0.3	0.2

Table 5: Ten clusters of size 50 and 100

Variable	Cox Regression	Gamma Frailty (Weibull)	Gamma Frailty (Exponential)	INLA
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
10 Clusters (n=50 per cluster)				
Gender				
Male	0.8825 (0.2154)	0.8820 (0.2101)	0.8803 (0.2102)	0.8755 (0.2099)
Female	Ref	Ref	Ref	Ref
Employment Status				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.9423 (0.3420)	0.9584 (0.3364)	0.9557 (0.3365)	0.9592 (0.3369)
CD4 Categories				
0-200	Ref	Ref	Ref	Ref
201-350	-.3801 (0.2576)	-.3890 (0.2527)	-.3877 (0.2527)	-.3889 (0.2521)
> 350	-.7845 (0.2731)	-.7984 (0.2685)	-.7974 (0.2686)	-.7986 (0.2681)
-2 Log L	660.7	651.7	652.8	N/A
DIC	N/A	N/A	N/A	289.2
Frailty Variance	N/A	0.3	0.3	0.2
10 Clusters (n=100 per cluster)				
Gender				
Male	0.8855 (0.1487)	0.8866 (0.1466)	0.8875 (0.1466)	0.8856 (0.1469)
Female	Ref	Ref	Ref	Ref
Employment Status				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.9215 (0.2337)	0.9283 (0.2320)	0.9286 (0.2320)	0.9286 (0.2320)
CD4 Categories				
0-200	Ref	Ref	Ref	Ref
201-350	-.3943 (0.1786)	-.3983 (0.1769)	-.3985 (0.1769)	-.3983 (0.1768)
> 350	-.7706 (0.1883)	-.7773 (0.1866)	-.7774 (0.1866)	-.7776 (0.1866)
-2 Log L	1600.5	-1300	-1301	N/A
DIC	N/A	N/A	N/A	525.4
Frailty Variance	N/A	0.3	0.3	0.3

Table 6: Twenty clusters of size 50 and 100

Variable	Cox Regression	Gamma Frailty (Weibull)	Gamma Frailty (Exponential)	INLA
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
20 Clusters (n=50 per cluster)				
Gender				
Male	0.8868 (0.1502)	0.8864 (0.1466)	0.8863 (0.1466)	0.8854 (0.1469)
Female	Ref	Ref	Ref	Ref
Employment Status				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.9078 (0.2338)	0.9230 (0.2307)	0.9233 (0.2307)	0.9231 (0.2307)
CD4 Categories				
0-200	Ref	Ref	Ref	Ref
201-350	-.3848 (0.1795)	-.3938 (0.1762)	-.3939 (0.1762)	-.3940 (0.1762)
> 350	-.7676 (0.1891)	-.7815 (0.1860)	-.7810 (0.1860)	-.7819 (0.1860)
-2 Log L	1330.3	1314	1315	N/A
DIC	N/A	N/A	N/A	570.6
Frailty Variance	N/A	0.4	0.4	0.3
20 Clusters (n=100 per cluster)				
Gender				
Male	0.8919 (0.1041)	0.8923 (0.1027)	0.8928 (0.1026)	0.8926 (0.1028)
Female	Ref	Ref	Ref	Ref
Employment Status				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.8992 (0.1618)	0.9057 (0.1607)	0.9073 (0.1607)	0.9069 (0.1607)
CD4 Categories				
0-200	Ref	Ref	Ref	Ref
201-350	-.3933 (0.1250)	-.3946 (0.1239)	-.3980 (0.1238)	-.3977 (0.1238)
> 350	-.7623 (0.1313)	-.7706 (0.1303)	-.7704 (0.1302)	-.7699 (0.1302)
-2 Log L	3215.4	2609	2613	N/A
DIC	N/A	N/A	N/A	1041.3
Frailty Variance	N/A	0.4	0.4	0.3

Table 7: Probability coverage

Variable	5 clusters (n=50)	5 Clusters (n=100)	10 Clusters (n=50)	10 Clusters (n=100)	20 Clusters (n=50)	20 Clusters (n=100)
Gender	94.8%	94.5%	94.7%	94.6%	94.7%	94.7%
Employment Status	94.7%	94.5%	94.9%	94.6%	95.0%	94.9%
CD4 Categories	95.3%	95.0%	94.9%	94.8%	94.9%	95.0%

Table 8: Mean squared error of parameter estimates

	10 Clusters, (n=50)	10 Clusters, (n=100)	20 Clusters, (n=50)	20 Clusters, (n=100)
Male				
Cox	0.04908	0.02211	0.02264	0.01065
INLA	**0.04673**	0.02142	**0.02131**	**0.01030**
Gamma Frailty (Weibull Baseline)	0.04679	0.02142	0.02138	0.01057
Gamma Frailty (Exponential Baseline)	0.04700	0.02141	0.02211	0.01103
Unemployment				
Cox	0.12115	0.05500	0.05725	0.02617
INLA	0.11756	0.05376	**0.05603**	0.02579
Gamma Frailty (Weibull Baseline)	0.11718	0.05372	0.05612	0.02575
Gamma Frailty (Exponential Baseline)	0.11493	0.05369	0.05619	0.02577
CD4 Count 201-350 mm^3				
Cox	0.06763	0.03068	0.03266	0.01612
INLA	**0.06505**	0.03058	**0.03148**	**0.01588**
Gamma Frailty (Weibull Baseline)	0.06537	0.03056	0.03152	0.01644
Gamma Frailty (Exponential Baseline)	0.06518	0.03060	0.03155	0.01592
CD4 Count > 350 mm^3				
Cox	0.08295	0.03640	0.03847	0.01693
INLA	0.08242	**0.03568**	**0.03740**	**0.01666**
Gamma Frailty (Weibull Baseline)	0.08172	0.03571	0.03747	0.01703
Gamma Frailty (Exponential Baseline)	0.08253	0.03567	0.03807	0.01668

Table 9: Comparison of models using HIV data

Variable	Cox-PH Estimate (SE)	Gamma Frailty with Weibull Hazard Estimate (SE)	Gamma Frailty with Exponential Hazard Estimate (SE)	INLA Mean (SD)
β_{Gender}				
Male	0.3006 (0.0906)	0.2844 (0.0908)	0.4485 (0.0913)	0.2849 (0.0908)
Female	Ref	Ref	Ref	Ref
$\beta_{Employment}$				
Employed	Ref	Ref	Ref	Ref
Unemployed	0.3598 (0.1122)	0.4011 (0.1130)	0.5954 (0.1119)	0.4025 (0.1130)
$\beta_{CD4\ Categories}$				
[0-200)	Ref	Ref	Ref	Ref
[200-350)	-0.8973 (0.1125)	-1.0016 (0.1136)	-1.2491 (0.1156)	-1.0031 (0.1137)
[350-500)	-1.3544 (0.1610)	-1.4539 (0.1621)	-1.5537 (0.1642)	-1.4559 (0.1622)
≥500	-1.6926 (0.1947)	-1.7767 (0.1956)	-1.9338 (0.1962)	-1.7789 (0.1956)
$\beta_{WHO\ Stage}$				
Stage 1 or 2	Ref	Ref	Ref	Ref
Stage 3 or 4	0.6225 (0.0899)	0.6156 (0.0920)	0.3806 (0.0976)	0.6176 (0.0920)
Fit statistics				
DIC	-	-	-	4615.66
-2 Log Likelihood	9387.44	5715.32	7593.3	-
Effective number of parameters	-	-	-	8.97
Frailty variance	-	0.5038	2.22	0.40=1/2.53

4.2.2.2 Application to HIV data

The Adult Wellness study enrolled 6,690 HIV-infected participants of whom three-quarters were females (n=5035). A total of 566 (8.5%) participants died during follow-up with the proportion of male deaths (11.4%, 189/1655) being higher than that of females (7.5%, 377/5035).

Findings comparing the four models as applied to data from an HIV-infected cohort from South Africa are presented in table 9. The estimates from the shared Gamma frailty[5] with exponential baseline hazard differed from those obtained using the other models. For example the estimate (se) for gender was 0.4485 (0.0913) while it was 0.3006 (0.0906), 0.2844 (0.0908) and 0.2849 (0.0908) for Cox-PH, Gamma frailty (Weibull baseline hazard) and INLA respectively. The Cox-PH model had the highest fit statistic value (-2 loglikelhood=9387.44) relative to other models. Within the Gamma frailty models, the model with a Weibull baseline hazard (-2 loglikelhood=5715.32) had a lower fit statistic relative to that with an exponential baseline hazard (-2 loglikelhood=7593.3). Within the frailty models, the estimates (standard errors) and means (standard deviations) of the Gamma frailty (with Weibull baseline hazard) and INLA respectively were similar.

The standard errors and standard deviation values were similar to 1 decimal place across all models. Parameter estimates were generally similar except for the Gamma frailty with exponential baseline hazard. A summary of the findings are presented in Appendix 2.

4.2.3 Semiparametric frailty modelling

In the comparison of the semiparametric gamma frailty models of 10 clusters of size 50 and 100, four maximisation procedures (penalized full and partial likelihood, INLA and H-likelihood) were compared (Figure 3a and 3b). Findings were close in the

measure of parameter estimates but INLA performed better. Findings were also similar under the lognormal frailty. Variability in general seemed higher in the smaller cluster of size 50 relative to the larger cluster of size 100.

Figure 4 presents the graphical plots showing the distribution of the standard errors as well as standard deviations from INLA. The MSE results (table 10) derived from the estimates of the parameters were also similar but INLA had consistently smaller values. MSE values determined from the larger cluster of size 100 had less variability compared to that derived from the cluster of size 50.

Gamma frailty variances from the four likelihood procedures were similar (table 11). INLA associated frailty variances that are associated with the piecewise log-constant baseline hazard were lower for both Gamma and lognormal frailties. On the contrary, those associated with the H-likelihood and penalized full likelihood were higher for both Gamma and lognormal frailties.

Efficiency of the parameter estimates is presented in table 12.

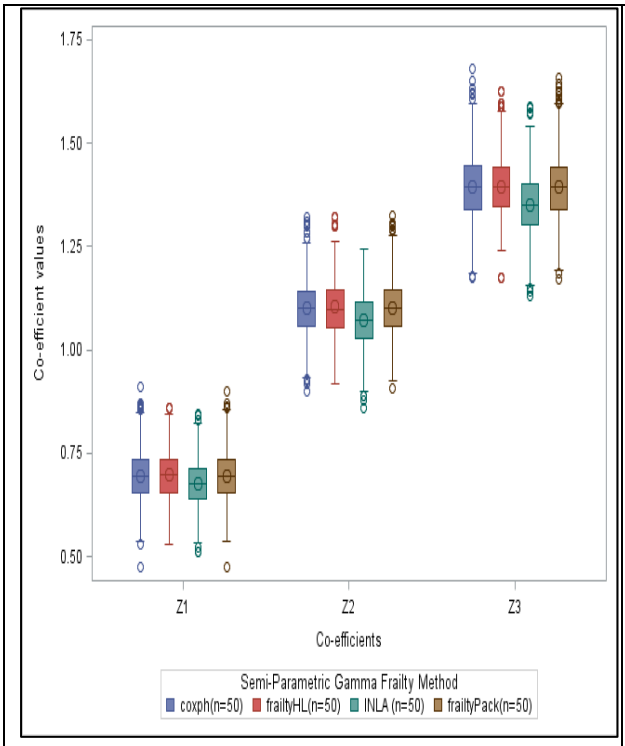


Figure 3a: Gamma frailty parameter estimates (n=50)

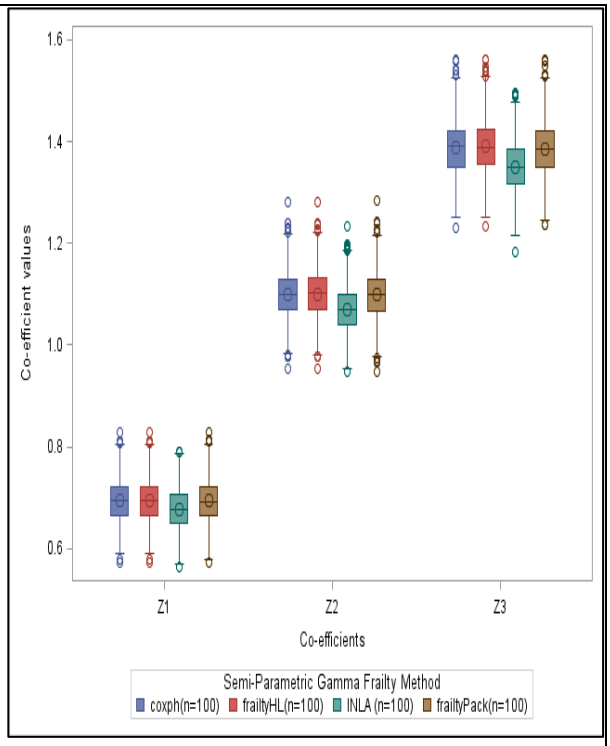


Figure 3b: Gamma frailty parameter estimates (n=100)

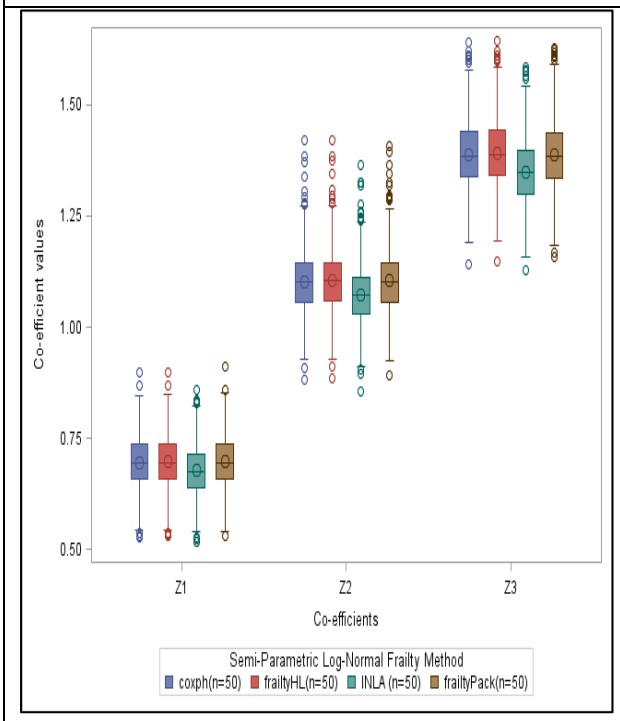


Figure 3c: Log-Normal frailty parameter estimates (n=50)

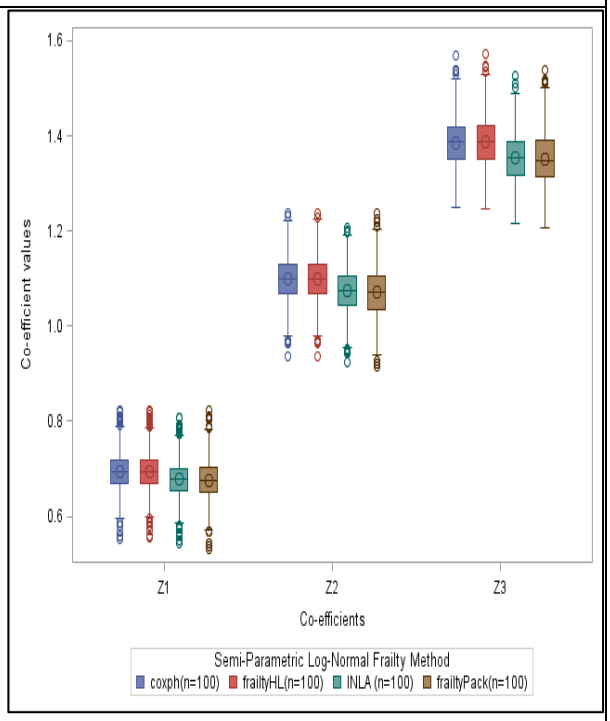


Figure 3d: Log-Normal frailty parameter estimates (n=100)

NB: n=50 and n=100 are the cluster sizes

Figure 3: Distribution of parameter estimates of the gamma and lognormal frailties

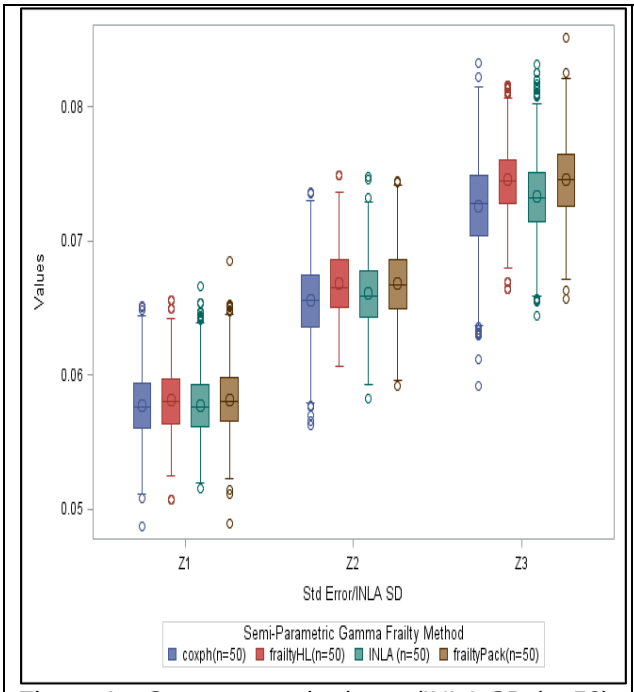


Figure 4a: Gamma standard error/INLA SD (n=50)

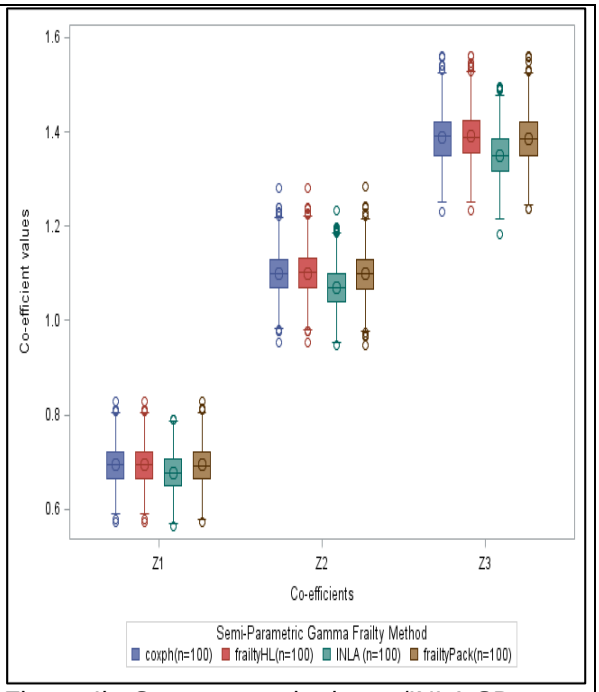


Figure 4b: Gamma standard error/INLA SD (n=100)

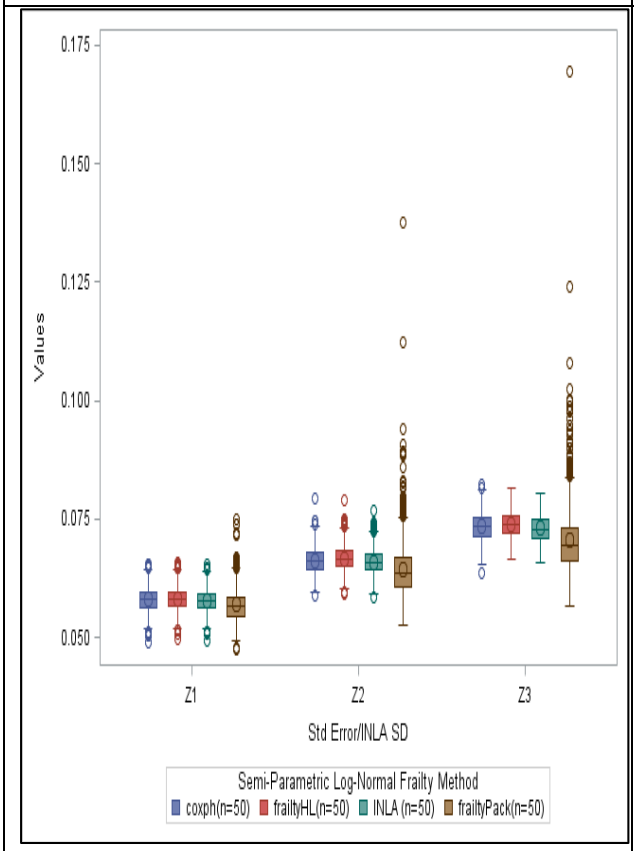


Figure 4c: Lognormal standard error/INLA SD (n=50)

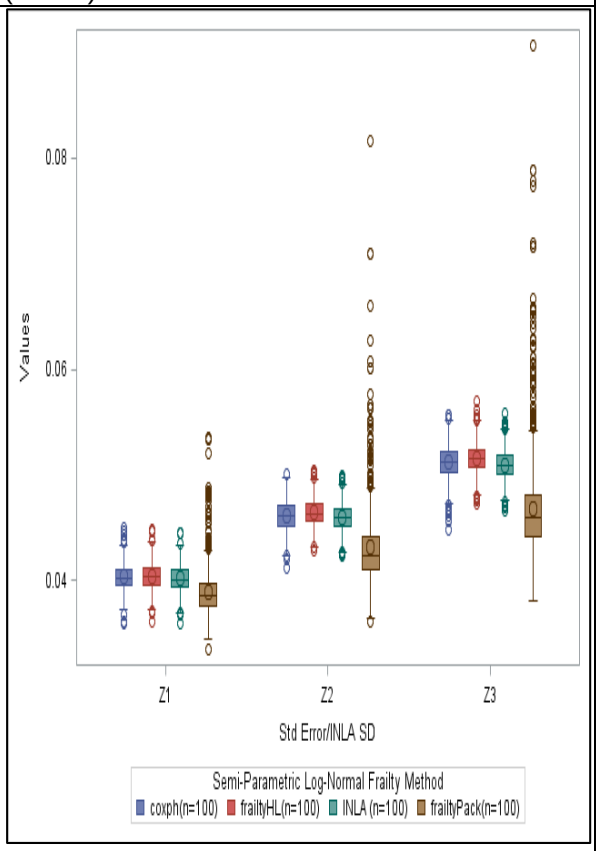


Figure 4d: Lognormal standard error/INLA SD (n=100)

NB: n=50 and n=100 are the cluster sizes

Figure 4: Distribution of standard errors/standard deviation

Table 10: Distribution of parameter estimate MSE

Parameter Estimates	INLA	frailtyHL	frailtyPack	coxph
Semiparametric Gamma Frailty				
Cluster size (n=50)				
Z1	0.0033	0.0036	0.0035	0.0035
Z2	0.0041	0.0046	0.0044	0.0044
Z3	0.0051	0.0054	0.0056	0.0058
Cluster size (n=100)				
Z1	0.0015	0.0017	0.0017	0.0017
Z2	0.0020	0.0022	0.0022	0.0022
Z3	0.0025	0.0026	0.0027	0.0027
Semiparametric Log-Normal Frailty				
Cluster size (n=50)				
Z1	0.0030	0.0033	0.0032	0.0032
Z2	0.0043	0.0047	0.0049	0.0046
Z3	0.0053	0.0058	0.0058	0.0058
Cluster size (n=100)				
Z1	0.0015	0.0016	0.0018	0.0016
Z2	0.0020	0.0022	0.0025	0.0022
Z3	0.0024	0.0026	0.0031	0.0026

NB: INLA- *Piecewise log-constant baseline hazard*; **frailtyHL**-*H-Likelihood*; **frailtyPack**-*Penalized full likelihood;* **CoxPH**-*Penalized partial likelihood*; *n=50 and n=100 are the cluster sizes*

Table 11: Median (IQR) distribution of the variances of the frailty distributions

Variable	INLA	FrailtyHL	FrailtyPack	CoxPh
Gamma frailty				
Gamma (Cluster size n=50) frailty variance	1.1 (1.0-1.1)	1.9 (1.4-2.4)	1.8 (1.3-2.3)	1.8 (1.3-2.3)
Gamma (Cluster size n=100) frailty variance	1.1 (1.0-1.1)	1.9 (1.4-2.5)	1.8 (1.3-2.3)	1.8 (1.3-2.3)
Log-Normal frailty				
Log-Normal (Cluster size n=50) frailty variance	1.1 (1.1-1.2)	1.9 (1.3-2.5)	1.8 (1.2-2.5)	1.2 (0.9-1.6)
Log-Normal (Cluster size n=100) frailty variance	1.1 (1.1-1.2)	1.8 (1.3-2.5)	1.7 (1.2-2.3)	1.2 (1.0-1.6)

NB: INLA- *Piecewise log-constant baseline hazard*; **frailtyHL**-*H-Likelihood*; **frailtyPack**-*Penalized full likelihood;* **CoxPH**-*Penalized partial likelihood*;

Table 12: Efficiency of parameter estimates

Method/package	Var ($\widehat{z_2}$)/Var ($\widehat{z_1}$)	Var ($\widehat{z_{23}}$)/Var ($\widehat{z_1}$)	Var ($\widehat{z_3}$)/Var($\widehat{z_2}$)
Semiparametric Gamma Frailty/Cluster size (n=50)			
INLA (n=50)	1.2499	1.5489	1.2392
frailtyHL (n=50)	1.2715	1.4880	1.1703
frailtyPack (n=50)	1.2698	1.6149	1.2717
frailtySurv (n=50)	1.2893	1.5493	1.2016
Coxph (n=50)	1.2573	1.6336	1.2993
Semiparametric Gamma Frailty/Cluster size (n=100)			
INLA (n=100)	1.2892	1.6207	1.2571
frailtyHL (n=100)	1.3026	1.6009	1.2289
frailtyPack (n=100)	1.3121	1.6184	1.2334
frailtySurv (n=100)	1.3282	1.6442	1.2380
Coxph (n=100)	1.3075	1.6156	1.2357
Semiparametric Log-Normal Frailty/Cluster size (n=50)			
INLA (n=50)	1.4303	1.7521	1.2250
frailtyHL (n=50)	1.4297	1.7731	1.2402
frailtyPack (n=50)	1.4955	1.7750	1.1869
Coxph (n=50)	1.4271	1.7763	1.2447
Semiparametric Log-Normal Frailty/Cluster size (n=100)			
INLA (n=100)	1.3762	1.6076	1.1681
frailtyHL (n=100)	1.3835	1.6242	1.1740
frailtyPack (n=100)	1.4163	1.7159	1.2115
Coxph (n=100)	1.3916	1.6197	1.1639

NB: Var =variance; z_1, z_2 and z_3 are parameter estimates; *n=50 and n=100 are the cluster sizes*

4.3 Mortality risk factors by setting in South Africa

4.3.1 Enrolment characteristics stratified by rural and urban setting

There were 6,690 enrolled participants (with 67% from PHRU). Relative to males, females were more at both sites (n=1,555 (70%) in rural; n=3,480 (78%) in urban). Rural participants were older (median age 36 vs 33 years) and had lower CD4 count (178 vs 303 cells/mm^3) whereas urban participants were associated with a higher BMI (24.9 vs. 21). Table 13 provides a further illustration of the characteristics during study entry.

4.3.2 Most recent socio-economic and clinical characteristics

Of those enrolled in the urban site, 63% (n=2,826) were not on HAART; only 30% (n=666) were not on HAART in the rural site. While the proportion of unemployed people was high at both sites, there were more males that were employed relative to females. Those enrolled in the rural site were more likely to have had TB relative to those in the urban (995/2221, 45% vs. 1134/4469, 25%; $p<0.0001$). Table 14 provides clinical and behavioural characteristics stratified by setting.

Table 13: Participant characteristics at enrolment by site

Variable	Rural (Tintswalo)			Urban (PHRU)		
	Overall	Male Gender	Female Gender	Overall	Male Gender	Female Gender
No. enrolled	2,221	666	1,555	4,469	989	3,480
Age (Median/IQR))	36.4 (31.0-44.1)	39.2 (33.3-48.1)	35.2 (30.0-42.5)	32.7 (28.2-38.1)	35.7 (31.5-41.5)	31.7 (27.4-37.0)
Age-Group in years						
[18-25] (%)	216 (9.7)	30 (4.5)	186 (12.0)	672 (15.0)	65 (6.6)	607 (17.4)
[26-30] (%)	342 (15.4)	72 (10.8)	270 (17.4)	1117 (25.0)	158 (16.0)	959 (27.6)
[31-34] (%)	420 (18.9)	111 (16.7)	309 (19.9)	984 (22.0)	222 (22.4)	762 (21.9)
[35-39] (%)	424 (19.1)	136 (20.4)	288 (18.5)	864 (19.3)	250 (25.3)	614 (17.6)
>39 (%)	819 (36.9)	317 (47.6)	502 (32.3)	832 (18.6)	294 (29.7)	538 (15.5)
CD4 Count						
(Median/IQR)	178.0 (81.0-333.6)	162.0 (69.0-289.9)	184.0 (86.0-350.6)	303.0 (156.0-470.7)	281.0 (142.0-443.7)	309.0 (163.0-478.2)
CD4 Categories						
[0-200] (%)	1235 (55.6)	398 (59.8)	837 (53.8)	1474 (33.0)	367 (37.1)	1107 (31.8)
[201-350] (%)	472 (21.3)	144 (21.6)	328 (21.1)	1102 (24.7)	241 (24.4)	861 (24.7)
[351-500] (%)	284 (12.8)	75 (11.3)	209 (13.4)	915 (20.5)	190 (19.2)	725 (20.8)
≥501 (%)	230 (10.4)	49 (7.4)	181 (11.6)	978 (21.9)	191 (19.3)	787 (22.6)

BMI (Median/IQR)	21.0 (17.9-25.0)	19.9 (17.2-23.4)	21.4 (18.4-25.8)	24.9 (20.7-29.2)	22.5 (18.2-26.4)	25.6 (21.4-30.0)
BMI						
Underweight (%)	638 (29.0)	241 (36.9)	397 (25.7)	594 (13.4)	240 (24.8)	354 (10.2)
Normal (%)	987 (44.9)	299 (45.7)	688 (44.6)	1612 (36.4)	393 (40.6)	1219 (35.2)
Obese/Overweight (%)	572 (26.0)	114 (17.4)	458 (29.7)	2226 (50.2)	336 (34.7)	1890 (54.6)
Crowding index (Median/IQR)	1.5 (1.0-2.3)	1.3 (0.8-2.0)	1.6 (1.0-2.5)	1.5 (1.0-2.3)	1.0 (1.0-2.0)	1.5 (1.0-2.5)
Household Income (IQR) in Rands	600.0 (189.8-1010.0)	780.0 (50.0-1500.0)	500.0 (190.0-974.7)	1080.0 (500.0-2130.0)	1200.0 (600.0-2400.0)	1020.0 (500.0-2020.0)
Household Income						
≤ R1,000 (%)	1516 (72.8)	423 (67.4)	1093 (75.2)	2161 (48.5)	453 (45.9)	1708 (49.2)
R1,001-R5,000 (%)	528 (25.4)	193 (30.7)	335 (23.1)	2092 (46.9)	487 (49.3)	1605 (46.3)
> R5,000 (%)	37 (1.8)	12 (1.9)	25 (1.7)	203 (4.6)	47 (4.8)	156 (4.5)

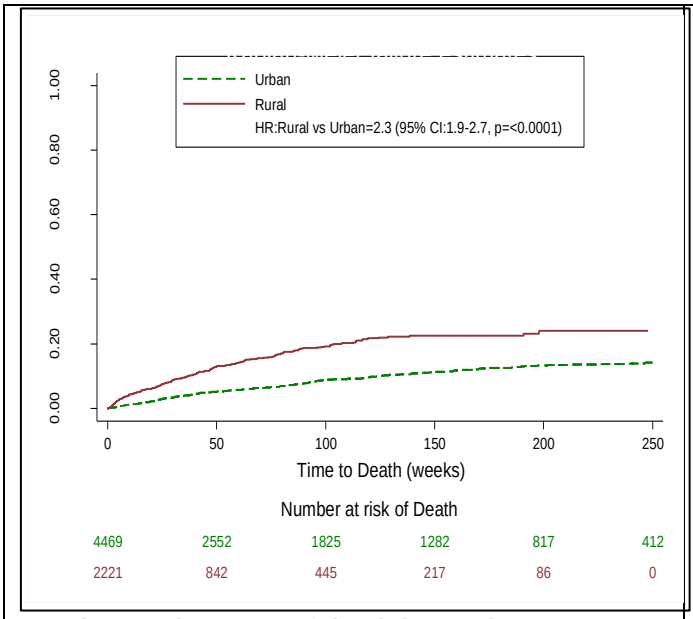
Table 14: Distribution of clinical and socio-behavioural factors

Variable	Rural		Urban	
	Male Gender (%)	Female Gender (%)	Male Gender (%)	Female Gender (%)
No. enrolled	666	1,555	989	3,480
Time on Treatment				
Not on Treatment				
(%)	174 (26)	492 (32)	555 (56)	2271 (65)
< 6 Months (%)	153 (23)	315 (20)	108 (11)	248 (7)
6-12 Months (%)	133 (20)	301 (19)	35 (4)	143 (4)
>12 Months (%)	206 (31)	447 (29)	291 (30)	818 (24)
Employment (ever)				
Yes (%)	143 (21)	153 (10)	400 (40)	909 (26)
No (%)	523 (79)	1402 (90)	589 (60)	2571 (74)
Smoking (ever)				
Yes (%)	295 (44)	63 (4)	706 (71)	578 (17)
No (%)	371 (56)	1492 (96)	283 (29)	2902 (83)
TB Infection (ever)				

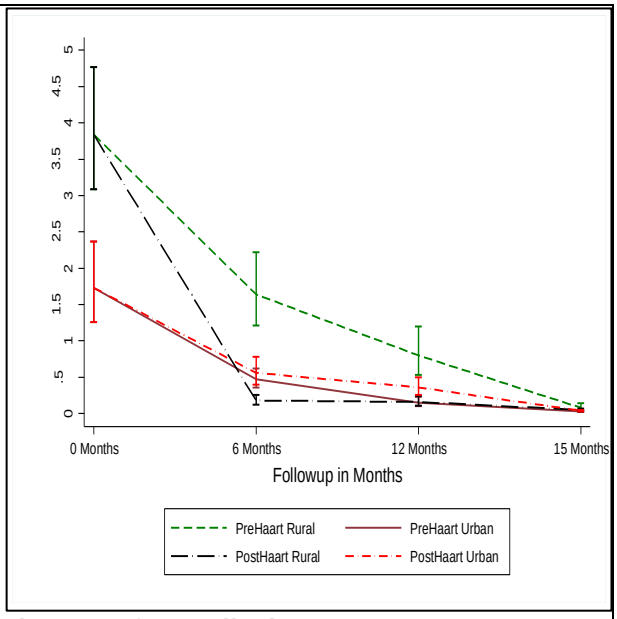
Yes (%)	350 (53)	645 (41)	313 (32)	821 (24)
No (%)	316 (47)	910 (59)	676 (68)	2659 (76)

4.3.3 Describing mortality by site

Approximately 11% and 7% of the rural and urban participants demised respectively during the course of the study. The rate of mortality was significantly elevated among rural participants in comparison to those recruited in the urban site. Their hazard of mortality was twice as high as that of the urban group. The overall rate of all-cause mortality as well as mortality on treatment is presented graphically in Figure 5 using the Kaplan-Meier plot.



5a: Time to the event of death by setting



5b: Rate of mortality by HAART status stratified by site

Figure 5: Survival analysis of death and its associated rate/100 person-years with or without treatment stratified by setting

Kaplan-Meier survival analysis curves for time-to-death overall as well as mortality rates per 100 person years with or without HAART stratified by setting are presented in Figures 5a and 5b. Similarly, Kaplan-Meier curves of time-to-death stratified by gender per site are presented in Figures 6a and 6b.

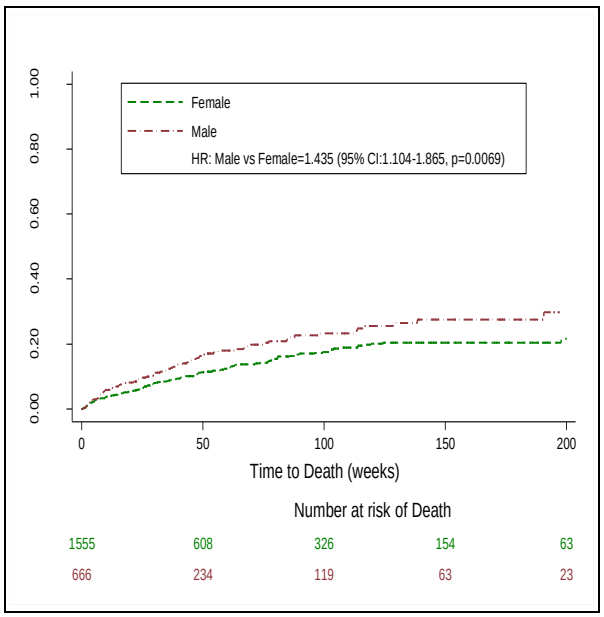


Figure 6a: Time to death by gender (rural)

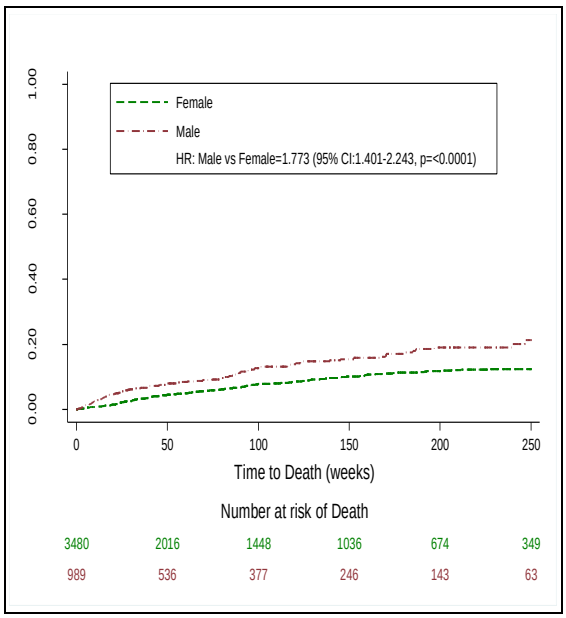


Figure 6b: Time to death by gender (urban)

Figure 6: Survival by gender in each setting

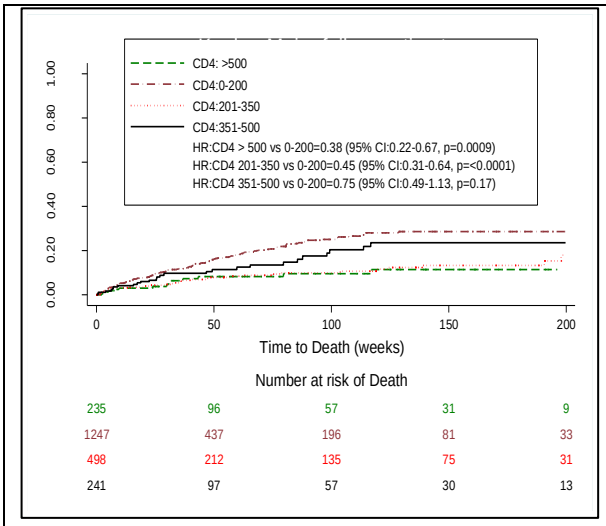


Figure 7a: Time to death by CD4 Count (rural)

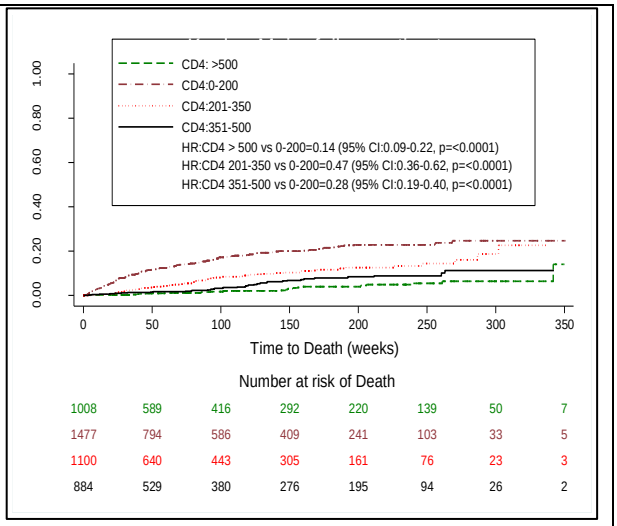


Figure 7b: Time to death by CD4 Count (urban)

Figure 7: Time to death by CD4 count within sites

In Figure 7, the Kaplan-Meier plot for time to death by CD4 count per site is presented. At both sites, those with CD4 count under 200 had a higher hazard for mortality relative to those with 201-350, 351-500 and > 500 cells/mm^3.

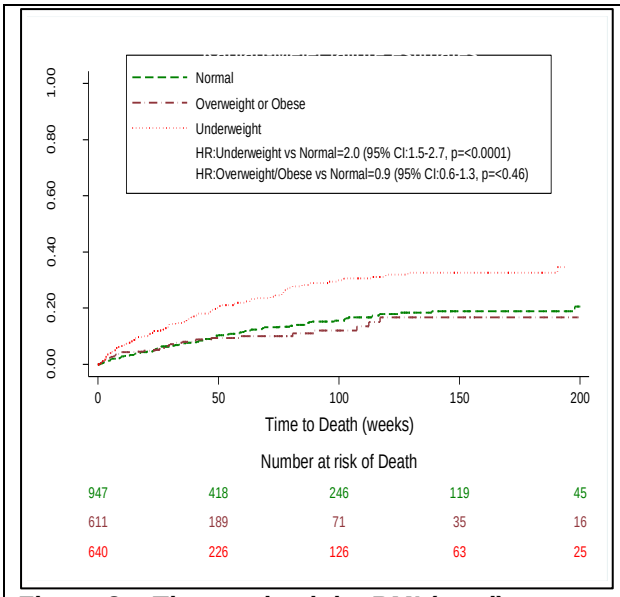


Figure 8a: Time to death by BMI (rural)

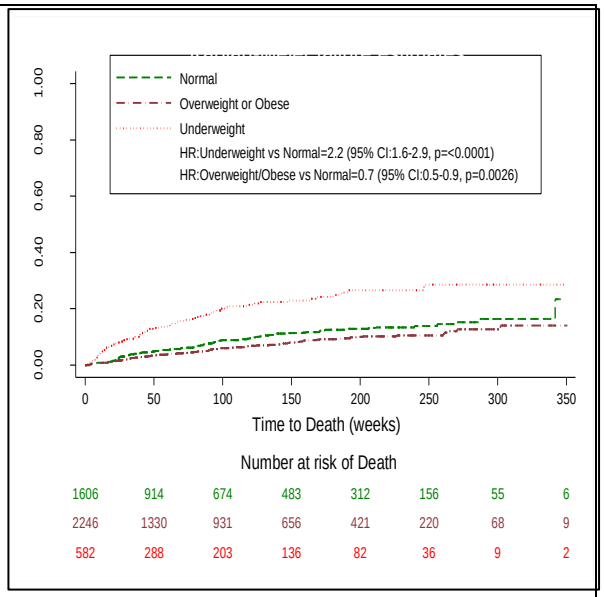


Figure 8b: Time to death by BMI (urban)

Figure 8: Time to death by BMI within sites

A higher risk of death at both sites (HR of 2 at both sites) was hazardous for underweight BMI whereas overweight/obese was protective in the urban group (Figure 8).

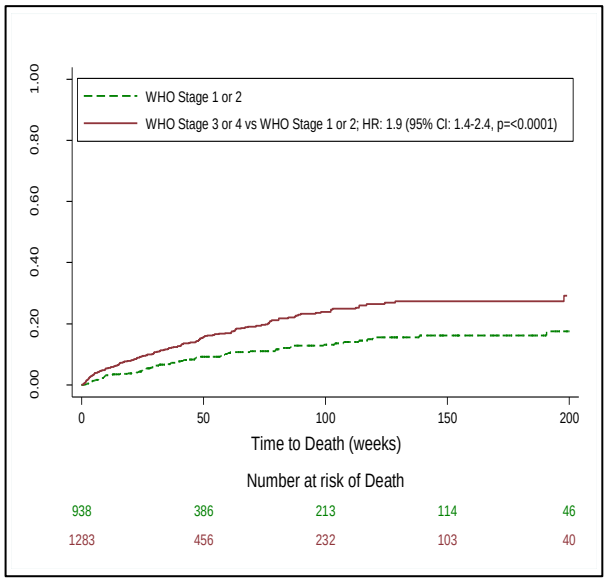


Figure 9a: Time to death by WHO Stage (rural)

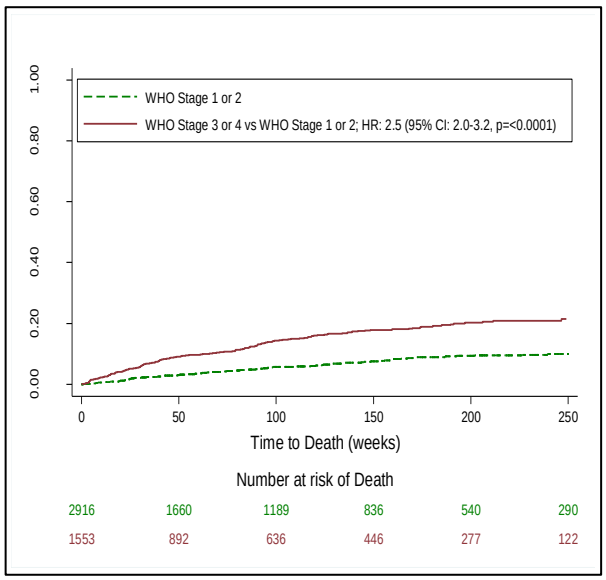


Figure 9b: Time to death by WHO Stage (urban)

Figure 9: Time to death by WHO Stage within sites

WHO stage 3 or 4 was hazardous elevating the risk of death relative to WHO stage 1 or 2 (Figure 9).

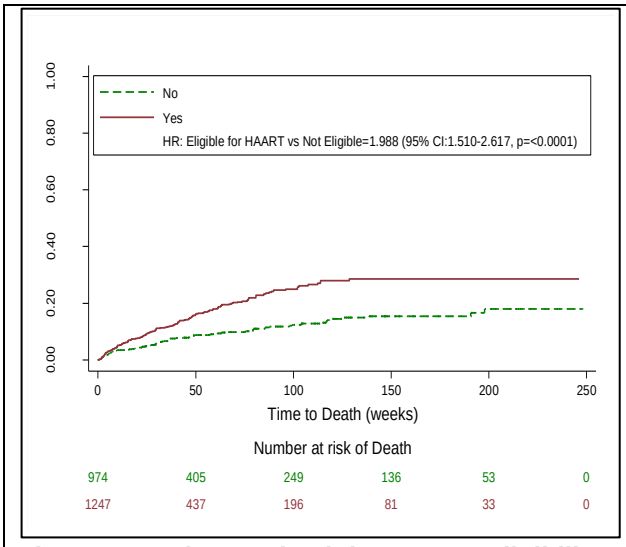


Figure 10a: Time to death by HAART eligibility (rural)

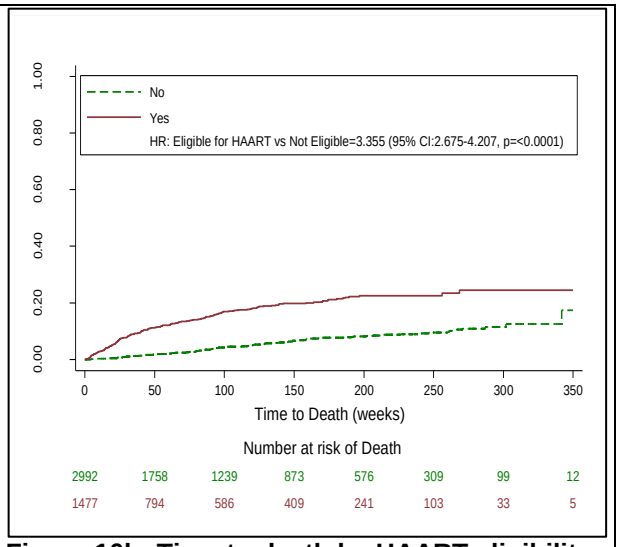


Figure 10b: Time to death by HAART eligibility (urban)

Figure 10: Time to death by eligibility (CD4 < 200 cells/µl) to start HAART by site

4.3.4 Mortality risk factors by setting

Being on treatment for 6-12 months (HR: 0.2, 95%CI: 0.1-0.4), > 12 months (HR: 0.1, 95%CI: 0.1-0.2) and a CD4 count > 200 cells/mm^3 (*CD4 201-350* HR: 0.3, 95%CI:0.1-0.5 ;*CD4 Count 351-500* HR: 0.4, 95%CI: 0.2-0.9; *CD4 Count >500* HR: 0.2, 95%CI: 0.1-0.5) reduced the hazard of death in rural participants. Those with underweight BMI had an elevated risk of death (table 15).

Treatment for >12 months (HR: 0.4, 95%CI: 0.3-0.5), a BMI value classified as overweight/obese (HR: 0.7, 95%CI: 0.6-0.9) and a CD4 count > 200 cells/mm^3 (*CD4 201-350* HR: 0.2, 95%CI:0.2-0.3 ;*CD4 Count 351-500* HR: 0.1, 95%CI: 0.1-0.2; *CD4 Count >500* HR: 0.1, 95%CI: 0.03-0.1) was associated with a lower hazard of death in the urban group. It was higher in those on treatment < 6 and 6-12 months, had underweight BMI, were unemployed, had ever smoked and got TB infection (table 15).

Both frailty variances were statistically significant suggesting the need to account for the effect of unobserved variables (heterogeneity) in the modelling process.

Table 15: Mortality risk factors by setting

Variable	Rural		Urban	
	Univariate	Multivariate	Univariate	Multivariate
	HR (95%CI)	HR (95%CI)	HR (95%CI)	HR (95%CI)
Gender				
Female	0.7 (0.5-1.1)	0.8 (0.5-1.3)	0.5 (0.4-0.7)	1.1 (0.8-1.4)
Male	Ref	Ref	Ref	Ref
Length of time on treatment				
<6 months	1.1 (0.6-2.2)	1.2 (0.6-2.2)	17.3 (11.7-24.8)	4.5 (3.0-6.6)
6-12 months	0.2 (0.1-0.4)	0.2 (0.1-0.4)	7.2 (4.9-10.4)	2.4 (1.6-3.5)
>12 months	0.1 (0.1-0.2)	0.1 (0.1-0.2)	1.0 (0.8-1.3)	0.4 (0.3-0.5)
No HAART	Ref	Ref	Ref	Ref
BMI Groups				
Underweight	2.2 (1.4-3.3)	1.7 (1.1-2.9)	4.3 (3.3-5.6)	2.5 (1.9-3.1)
Overweight/Obese	0.6 (0.3-0.98)	0.6 (0.3-1.2)	0.6 (0.4-0.7)	0.7 (0.6-0.90)
Normal	Ref	Ref	Ref	Ref
CD4 Count (cells/mm^3)				
201-350	0.3 (0.1-0.6)	0.3 (0.1-0.5)	0.2 (0.1-0.2)	0.2 (0.2-0.3)

351-500	0.5 (0.2-1.1)	0.4 (0.2-0.9)	0.1 (0.1-0.1)	0.1 (0.1-0.2)
>500	0.3 (0.1-0.7)	0.2 (0.1-0.5)	0.04 (0.02-0.1)	0.1 (0.03-0.1)
0-200	Ref	Ref	Ref	Ref
Employment (ever)				
No	0.95 (0.5-1.8)	1.1 (0.6-2.2)	2.0 (1.6-2.7)	1.9 (1.4-2.5)
Yes	Ref	Ref	Ref	Ref
Smoking (ever)				
Yes	0.99 (0.6-1.6)	-	2.1 (1.7-2.6)	1.4 (1.04-1.8)
No	Ref		Ref	Ref
TB infection (ever)				
Yes	1.7 (1.1-2.5)	1.6 (0.98-2.5)	2.8 (2.3-3.5)	2.8 (2.2-3.5)
No	Ref	Ref	Ref	Ref

4.4 Perceptions of risk factors for mortality-qualitative study

4.4.1 Demographics

Of the 20 participants attending the PHRU clinic for VCT counselling services that were enrolled, 17 were females. Their ages ranged from 18 to 63 years, highest school grade completed was between 2nd and 12th Grades and 9 reported Isizulu as their main language. They reported living in households with between 1 and 9 people and the total rooms in their houses ranged from 1 to 17. There were 10 and 7 participants from female headed households where the head of household was aged 18-60 and > 60 years respectively. Participants reported knowing the deceased persons for between 5 and 40 years. Eight participants were employed, 8 were unemployed and 4 were students in tertiary institutions. Deceased persons were either an aunt (n=5), cousin (n=7), son (n=2) or other relative (n=6).

Participants reported that the deceased persons were aged between 18 and 54 years, 8 reported IsiZulu as their main language and their highest school grade was between 2 and 12. Of the 20 deceased persons, 15 were females. Among the deceased persons, 12 were employed, 6 were unemployed and 2 were students in tertiary institutions.

4.4.2 Emerging themes on perceptions of mortality risk factors

Following the interviews, several themes emerged that were perceived as having an association with mortality among the deceased persons. Participants perceived that engagement in risky sexual behaviour including multiple sexual partnerships may have a role in worsening the situation for an HIV-infected person. A negative attitude by healthcare workers towards HIV-infected people was perceived to be associated with avoidance in seeking care. Those who believed in the healing power of religion and ignored antiretroviral treatment in favour of traditional medicine were perceived to have had their health deteriorate. Food security such as availability of regular meals as well as a good social support structure such as receiving help in domestic chores was perceived as important in keeping the HIV-infected persons healthy.

4.4.2.1 Perceived engagement in risky sexual behaviour

Participants reported about the deceased's risky behaviour such as multiple partnerships, unprotected sex and alcohol use. Some of the deceased were perceived to be involved in multiple relationships whilst being married. For instance, a participant stated the following of her father:

> *"I stayed with him for that 6 years and he was alright but you could see that even when he was here in Joburg, he still had a gift of liking women…he used to take me with him when he went to see his women (girlfriends). He was aware that he is with me. He had 4 (wives)." (Female Participant 1 in relation to her deceased father)*

Another provided the following summary of her friend's husband:

> *"She told me about (her husband) from what he does and so on. She said (her husband) would come back home in the morning and he knew very well that we (husband and deceased) are sick but he still did the things he did. That was reason why I moved out of the house cause he made me sick. He should have done the right thing. I joked about it and told her that she shouldn't have had just one boyfriend in life till they got married. She (deceased) said that is one thing that hurt her and made her sad. On the other hand, her partner (husband of the deceased) went on and slept with multiple partners" (Female participant 6 in relation to her deceased friend's husband)*

Participants commonly reported that once the deceased persons had improved while on treatment, they defaulted on their treatment. Thereafter they went back to their previous habit of drinking alcohol and engaging in multiple partnerships. A participant stated the following of her aunt:

“Yes she (deceased person) recovered then she stopped taking the treatment and went back to drinking again, continued with her boyfriends and that’s where she fell ill and got worse.” (Female participant 12 in relation to her deceased aunt)

Participants reported that engagement in unprotected sex occurred and it was unclear if disclosure of their HIV status happened. A female participant quoted the following of her aunt:

“*Now she (deceased person) is in a relationship with a pastor who has a wife and they don’t use a condom and who knows if (my aunt) told him that she is sick*” *(Female participant 6 in relation to her deceased aunt)*

It was perceived that the deceased persons engaged in multiple partnerships and unprotected sex while knowing their HIV status. Furthermore, it was unclear if they disclosed their HIV status to their partners. Defaulting on treatment was a more likely cause of death in those engaging in risky sexual behaviour.

4.4.2.2 Negative attitude by healthcare workers

A negative attitude by the healthcare workers was perceived to be a barrier in seeking care in public health facilities. The quote below summarises the advice given by a clinician to the deceased aunt of a participant:

“He (the doctor) told her to terminate (her pregnancy) but she didn’t want to and she never went back to the doctor and that is why she never went to the clinic to get medication to protect the baby from getting infected.” (Female participant 6 in relation to her aunt)

Another participant reported on the perceived negative attitude shown to her deceased relative by the nurses in the local clinic. The perceived negative attitude to a sister-in-law is summarised below:

> *"She (deceased person) said that nurses were not treating her well. I asked her why she didn't tell us because we could ask our aunt to take her to the (local) clinic and you (the deceased) would take the treatment there. She said that she was afraid. I asked her, how long would she be afraid when she had kids who were already complaining that their mother is always sleeping and they didn't understand what was going on." (Female participant 16 in relation to her sister-inlaw)*

Sometimes when participants reported on the perceived negative attitude by clinicians, they failed to be specific. A female participant provided the following summary of her deceased aunt's experience:

> ***"****The only thing she (deceased person) told me was that the doctors are rude. That is the only thing I knew. People usually used to say that. But she (deceased person) would never explain to you why she's saying they are rude. I would ask why you (deceased person) say they are rude and she would say they are rude. And the nurses when you call them, they take their time (to respond)" (Female participant 9 in relation to her deceased aunt****"***

It was perceived that the negative attitude by healthcare workers discouraged the deceased persons from seeking care in the public health care facilities and that may have led to deteriorating health.

4.4.2.3 Healing power of religion and traditional medicine

The use of religious prayers as well as traditional medicine for purposes of healing was espoused during the interviews. Others perceived witchcraft as the cause of their HIV infection. They believed someone bewitched them into their status.

The perceived use of religious prayers as a measure of healing by an aunt is summarised in the following quote by a female participant:

> *"She (deceased) believed in the church. She never went to sangomas (traditional healer). She went (to the church) because she loved it. She believed in it. She knew even if she was HIV (infected), God is there. She believed that miracles do happen" (Female Participant 6 in relation to her deceased aunt)*

Others attended churches where the clergy provided holy water or tea. Holy water refers to water that has been blessed by the clergy and apparently contains healing powers. Similarly holy tea is blessed tea. This is summarised in the following quotes by relatives of deceased persons:

> *"She (deceased person) believed in church. They asked the pastor and the congregation to come to our house. Then they prayed for her (deceased person). And they prayed for the (holy) water and told her how she must drink it and she must wash with it." (Female Participant 10 in relation to her deceased cousin)*

> *"He had a belief that since he is a member of (name of church) that he will get healed (by drinking tea). It was tea that they got from the church. How can you get healed by drinking tea? I even asked him (deceased person). Has he ever seen anyone get healed by drinking tea" (Male Participant 11 in relation to his deceased brother)*

The use of traditional medicine in managing HIV emerged during the interview. It was perceived that traditional medicine, provided by sangomas (traditional healers), was favoured by either the deceased person or their families. This is encapsulated by relatives of deceased persons in the following quotes:

> *"You know they even tried to consult a Sangoma to help her to get better. You know the traditional healers said we need to slaughter a chicken and make a ceremony for her. Oh and the Sangoma said she must drink the chicken's blood." (Female Participant 9 in relation to her deceased aunt)*

> *"They (family) did take him to a traditional healer and he was given traditional tonic and medication and things like that. I am not sure, but they took him to a traditional healer because of the chest pains suspecting that he was bewitched in his beer. You know how people are? So they suspected witchcraft because he was complaining of pain at the side of his chest." (Female Participant 20 in relation to his deceased uncle)*

Other participants reported that some family members perceived witchcraft as the cause of the illnesses suffered by the deceased. One participant summarised the following of her aunt:

> *"No I don't think she (deceased person) told them because they did not visit her while she was sick. My aunt suspected witchcraft and because I knew the situation, I told her not to point fingers because I was here when mother was very sick. I saw her. Some people are like that. They blame witches for everything. When things go wrong, when their children fail at school, they say it is witchcraft." (Female Participant 8 in relation to her deceased mother)*

The significance attached to religious prayers, traditional medicine and witchcraft emerged during the interview. These were perceived as avenues that offered alternative forms of care for HIV infected individuals.

4.4.2.4 Food security

During the interviews, it emerged that food security was perceived as a matter of concern for HIV infected individuals that had stopped working due to HIV associated opportunistic infections. Participants noted that availability of food was important in maintaining good health especially for those who are HIV-infected. Participants described the nature of symptoms of the deceased persons, their experiences in relation to quitting employment and its effect on food security. Additionally, the support of the family, relatives and friends signified the importance of the social support structure. A female participant described her mother's symptoms that led her to quit employment as follows:

> *"When she started coughing seriously, she had to quit (working). It was her decision. Her body became weak and she was not coping. She was very sick and feeling pain. She could not go to work." (Female Participant 8 in relation to her deceased mother)*

Unfortunately not working led to loss of income followed by reduced availability of food especially for those who were self-employed and running their own businesses. A participant stated the following of her deceased aunt:

> *"She used to sell food in a shop (in Soweto). So when she got sick, she couldn't go sell. So there was no one running her business. She was used to doing things on her own. That actually stressed her because on the other hand she didn't have money. So my uncle that lives in Pimville (a township in Soweto) and another aunt that lives with us helped her. They told her if she needs anything, they will help. They told her not to worry*

but to focus on her health. That is when she relaxed a bit." (Female Participant 7 in relation to her deceased aunt)

Some of the deceased persons were unemployed and relied on others such as parents, friends or partners for food. A female participant summarises the experience of her deceased cousin as follows:

"Her mother would bring her food but then she said there were times when she had nothing to eat. So she would have to go eat with her friends or her boyfriends, something like that." (Female Participant 2 in relation to her deceased cousin)

Not only did the deceased persons stop working and impacting on household food security but also those who cared for them. As one participant noted, of the mother of one of the deceased persons:

"Yes, then she (mother of deceased) stopped (working) when her daughter got sick. Yes, she (mother of deceased) stopped working so that she can take care of…(her daughter) and her child (grandchild)" (Female Participant 5 in relation to her deceased cousin)

Poor health of the deceased persons influenced their decision to stop working. Symptoms such as coughing, feeling weak and pain were commonly reported. Additionally, participants reported on the association between income and food security. Those who stopped working due to illness lost their income and livelihood. Participants noted that the support provided by the family, relatives and friends enabled the deceased persons an opportunity to live without worrying about food.

4.4.2.5 Social support structure

The perceived importance of the social support structure such as the family, relatives and friends emerged during the interviews. Discussions revealed that deceased persons were able to receive care at home including support in attending to domestic chores or attending clinics. While other deceased persons received family support, others did not. A female participant noted the following of her aunt who never got support from the boyfriend:

> *"The only thing that I heard was about her (aunty's) boyfriend. She (aunty) complained that since she started being sick, he (boyfriend) didn't come and see her. He wasn't there for her". (Female Participant 7 in relation to her deceased aunt)*

There were those who became weak, frail and bedridden and required support with domestic cores such as washing their dirty linen. As one female participant noted of her aunt, the frail required help.

> *"They did everything for her. Her father would wake up early in the morning to make porridge for her and her mother would give her a bath. She would change her linen for her. She was very clean. Every Sunday, they would get someone to drive her to church because she liked church. They would drive her everywhere she would want to go " (Female Participant 6 in relation to her deceased aunt)*

Participants reported that support to move around and take treatment on time was possible due to the availability of a social support structure. A participant summarised her thoughts in relation to her aunt as follows:

"She made food for her. She made sure that she takes her pills on time. If she needed her to help her stand, she would do that for her. The pills that she was taking were grandpas (analgesics) that she bought herself at the spaza shop (informal store). She would boil water for her and also rub her feet if they needed to get rubbed until she went to the hospital" (Female Participant 7 in relation to her deceased aunt)

The perceived importance of a social support structure to provide support to frail HIV-infected persons was emphasised during the interviews. The social support structure enabled the deceased HIV-infected persons to lead a comfortable life. Timeous initiation of treatment was possible where social support was available.

5.0 Discussion

This study was conducted to show the benefit of utilising a mixed methods approach incorporating both quantitative and qualitative analysis methods to complementarily. The methods were applied on the Adult Wellness Study involving HIV-infected individuals that was conducted at two sites in South Africa over seven years (2003-2010). The statistical techniques applied in the analysis of data assessing the risk of death, and their suitability were critically reviewed.[54] The assessment was conducted using a modified statistical classification method that has been used previously in other fields in health research.[66] Additionally, a selected number of time-to-event methods were compared using simulations with the aim of providing guidance on the most suitable technique to utilise in order to account for the effect of unobserved variables (heterogeneity), a factor that is often ignored. Risk factors for mortality in the Adult Wellness study were assessed by study site using INLA, a Bayesian survival frailty model.[58] The two sites were classified into rural (Tintswalo hospital in Mpumalanga) and urban (PHRU in Chris Hani Baragwanath hospital, Soweto). Thereafter, a qualitative study was conducted in the urban site (PHRU in Soweto) and analysed. The findings were subsequently used to complement those from the quantitative risk factors analysis.

This section discusses the results in the context of existing information. Findings are discussed under the following key areas: critical review of literature, comparison of time-to-event models, incorporation of latent variables and perceived risk factors.

5.1 Critical review of literature

This study critically reviewed articles on mortality risk factors in HIV-infected individuals in the periods: 2002-2006 and 2007-2011 to determine methods used and whether they improved over time. Basic inferential statistical techniques were likely to be reported in these two periods. Our findings concurred with others who also assessed commonly used statistical methods for such outcomes and study design.[35, 95] The use of frailty modelling that incorporates the effect of unobserved variables was more common in the later period of 2007-2011.[96] Though analysis

using sub-optimal methods has previously been reported,[97-101] there were few in our review; limited use of the advanced survival frailty modelling was observed. It may be that there is a compelling need for capacity building in the use of complex survival analysis techniques in health related studies. Additional findings from this study are presented in Appendix 1.

5.2 Selecting an appropriate time-to-event model

A selected number of parametric and semiparametric frailty models were compared to examine how well they incorporate unobserved (latent) measures when assessing the risk factors for mortality. Cox regression methods with shared frailty performed better than the classical Cox regression in the parametric simulation parameters. In general, the Bayesian INLA had a better fit under the MSE criteria. While parameter estimates were similar in the semiparametric models, the INLA models consistently led to a better fit using the MSE criteria. Our findings agree with previous work showing a better fit with lower MSE values in the semiparametric models. [2] Additionally, INLA, a Bayesian technique, performed better than other methods as previously shown.[8, 44, 102, 103] Appendix 2 presents further clarity of the methods used and findings.

5.3 Accounting for unobserved heterogeneity variables

Our findings suggest that where possible, it may be worthwhile to fit a model that adjusts for clusters such as clinics that act as surrogates for unobserved variables. Participants in the rural site had more heterogeneity compared to those from the urban site suggesting a higher risk for mortality. Opportunistic infections, socio-economic and socio-cultural factors are some of the potential unobserved variables that could be associated with mortality. Other publications that have determined factors associated with mortality found that unobserved variables had a significant effect on the outcome.[6, 15, 16] Ignoring unobserved variables may lead to biased conclusions that are likely to negatively affect the management and care of HIV-infected patients. This may lead to disease progression and death. Hence using several complementary

methods in a mixed methods framework may provide insights that will improve patient care. An in-depth presentation of findings from this study is reported in Appendix 3.

5.4 Perceived risk factors

Since unobserved variables were significantly associated with mortality in HIV-infected people, qualitative analysis was explored with the aim of identifying the potential unobserved variables and using them to complement the quantitative findings. Participants were required to have been bereaved due to HIV. Several themes emerged including risky behaviour, negative attitude by healthcare workers, religious prayers and traditional medicine, food security and social support structure.

Participants perceived an elevated mortality risk potentially associated with risky behaviour including multiple sexual partnerships and unprotected sex. The role of multiple sexual partnerships as well as unprotected sex in relation to HIV and new HIV cases has widely been researched.[104-109] Additionally, evidence shows that HIV infection is likely to occur in stable relationships.[105, 108] Unfortunately partners only become aware of their HIV status when their health condition deteriorated. Often this leads to late diagnosis and initiation of therapy when the immune system is severely immunosuppressed leading to early mortality.[110, 111] While antiretroviral therapy is free in public health facilities, mortality rates associated with HIV still remain high. Following the qualitative analysis findings on the perceived relationship between multiple partnerships and mortality, we postulate that defaulting on treatment was a more likely cause of death as indicated by one of the key informants.

Negative attitude by healthcare workers was perceived to have discouraged the deceased persons from seeking care in the public healthcare facilities. This finding is consistent with existing literature about discrimination of HIV-infected people by healthcare workers.[112-115] While there has been more awareness about management of HIV and HIV-infected patients among healthcare workers,[116] there is still a small proportion that continues to display negative attitude towards HIV-infected individuals. Discriminatory practice against HIV-infected individuals

dissuades them from seeking care in public healthcare facilities. Discrimination of HIV-infected individuals may lead to poor health due to low levels of treatment adherence and utilisation of healthcare services.[113, 116, 117]

In our findings, participants perceived that some of their deceased relations had more faith in healing through religious prayers. They believed that going to church, asking for forgiveness and praying might lead to a miracle that would reverse their HIV condition. This was reinforced by religious leaders willing to offer prayers and support towards healing. Furthermore, some clergy even presented a form of blessed water suggesting that consuming it would lead to healing from HIV. The use of religion as a coping mechanism for HIV has previously been documented.[118, 119] While religion frames the way of life for many people including HIV-infected individuals, the choice of ignoring care and initiation of therapy for those eligible to start treatment for the sake of religious prayers may be detrimental to their health. It may worsen the health of immunosuppressed individuals leading to negative consequences such as death.[110, 111] It may be that healthcare providers need to consider an integrated approach that provides care to HIV-infected individuals while at the same time accommodating their religious way of life such as allowing them to continue seeking healing through prayers. The interplay between religion and the belief that it leads to healing of HIV requires further research.

We found preference for traditional medicine over antiretroviral therapy by some of the deceased persons. The perceived benefits associated with the use of traditional medicine in treating HIV remains unclear. However, it may be that its use is associated with the need to identify with one's indigenous identity, family expectations, privacy and confidentiality.[120, 121] Furthermore use of traditional medicine in combination with antiretroviral therapy may lead to drug interactions[122] and poor health outcomes such as death. Healthcare providers should consider use of both traditional and contemporary medicine under their supervision in order to mitigate any adverse events and manage them accordingly.[120] This should be feasible within systems where regulations governing use of traditional medicine are in place such as in South Africa.

The combined effect of a HIV-related death and food insecurity has previously been studied where it has been shown that social and development implications cannot be ignored.[123, 124] Participants perceived a threat to food security in the livelihood of deceased persons due to loss of income when they ceased employment. This was often due to HIV-associated illnesses. Mortality is disproportionately high in the HIV-infected people who are often economically active.[123, 125, 126] As a result, the livelihood of their dependents is put at risk. Our findings are consistent with previous work linking food insecurity and mortality.[124, 127-129]

Our findings showed the importance of providing social support to HIV-infected individuals. Participants perceived that provision of social support such as helping in domestic chores, food security and support to move around may have reduced the burden of disease and potentially reduced the risk of mortality. While literature on social support such as home-based care for HIV-infected people is widely available,[130-134] there is limited information if any to show how provision of social support is linked to HIV-associated mortality.[135] The need for further research in low socio-economic settings on the interaction between social support and HIV-associated morbidity and mortality is required.

5.5 Linkage between study findings and conceptual framework

This study has confirmed the benefits of using a mixed methods approach to complement findings using the critical realism framework as presented in Figure 1 of section 1.5. A critical review of the literature was performed to determine existing gaps followed by a comparison of time-to-event methods to determine a model that would be appropriate in accounting for unobserved variables. Thereafter, the model was applied to real data from a South African HIV study conducted at two sites. Findings from the quantitative analysis were complemented by the qualitative study to provide further insights. The critical realism conceptual framework enabled triangulation of results increasing the scope of our findings. Observed and important unobserved covariates from both methods were identified as influencing mortality in HIV-infected people. Additionally, the framework increased our confidence in the hypothesis and

findings by providing a wider view. Following both approaches, risk factors for mortality are summarised as demographic, socio-economic, clinical, socio-cultural, behavioural, social support and attitude of healthcare workers. More research is required to understand how socio-cultural, behavioural and social support affects morbidity and mortality in HIV infected individuals.

5.6 Study limitations

5.6.1 Eligibility criteria

Publications included in the critical review paper on statistical methods were only those published in Pubmed/Medline. Additionally only publications written in the English language were considered. Therefore any other publications written in a foreign language and indexed in a database other than Pubmed/Medline were excluded. Additionally, our hypothesis testing may have been influenced by the non-random sample since this was dependent on the available prospective and retrospective studies found during our search criteria; the comparison of proportions may also have been influenced by the categories of years chosen.

5.6.2 Missing data

In the Adult Wellness data, some variables such as CD4 count had a low percentage of missing data. For such variables, multiple imputation technique was used assuming data were missing at random and this may have introduced bias.

5.6.3 Social desirability bias

In the HIV study conducted in South Africa, some variables were self-reported suggesting the possibility of social desirability bias, where the respondent provides responses that are favourably viewed.

5.6.4 Misclassification of data

Some variables such as “ever employed” were assessed cumulatively suggesting that misclassification may have occurred because they were not instantaneous. It is also likely that

some outcome measures captured as loss to follow-up may have later resulted into deaths that were not subsequently captured in our database.

5.6.5 Perceptions of participants

The qualitative study used participants from the urban site (PHRU), and their responses may not reflect those from the rural site. Additionally, participants were enrolled after the Adult Wellness HIV study had ended. It is likely that findings may have differed if the study had been implemented at both sites at the same time. Furthermore, the study period was different, with the qualitative study being conducted a few years after the end of the quantitative one, in which perceptions may have changed or participants experienced recall bias. Additionally, all of them were VCT clients (and thus likely to follow a line of thought introduced to them by the VCT counsellors).

5.6.7 HAART Adherence data

Use of HAART and its adherence is an important factor for positive outcomes. In the Adult Wellness study, adherence to HAART treatment was collected but not reliably. Hence it was not possible to assess the impact of HAART on patient clinical outcomes.

5.7 Strengths of the study

The use of simulated data provided a statistically powerful platform to compare the selected time-to-event methods and thereafter apply them to real data. Actual cohort data from a real study may be influenced by uncontrolled factors, introducing bias in the results. As such, they may not be the best data to use for comparing methods other than for confirmatory work.

An advantage of the HIV data from South Africa was that it was a large sample with long-term follow-up that allowed for the fitting of various survival frailty models, accounting for unobserved variables.

Analysis using mixed methods afforded a complementary approach allowing a wider scope of coverage in identifying important covariates. While the frailty model determined the significance of unobserved variables, the qualitative analysis was helpful in identifying them.

6.0 Conclusions, implications and future research

6.1 Conclusions and implications

The benefits associated with using a mixed methods approach were explored in this study in which the need for capacity building in the use of complex time-to-event methods was observed. These involve using statistical techniques that allow for modelling of time-to-event outcomes while at the same time accounting for unobserved measures.

This study presented results on the comparison of selected parametric and semiparametric time-to-event methods in which parameter estimates, standard errors and their frailty variances were assessed. In general, frailty models performed better with the Bayesian INLA performing best. While frailty models can be complex to implement, they reduce bias by incorporating the effect of unobserved covariates hence improving interpretation of study conclusions.[4, 5, 7, 44] Statistical practitioners need to ensure robustness in the analysis of data.

A comprehensive analysis of data is likely to provide stakeholders (such as policy makers or HIV care managers) a clear understanding of the existing gaps and types of interventions to implement. Interventions could be adapted to settings allowing for adaptability and scaling up both locally and regionally

6.2 Areas of further research

Following our findings, the following areas for further research are proposed:

1. In the analysis of qualitative data, participants perceived the importance of a good social support structure in mitigating the risk of mortality in HIV infected individuals. Our findings though showed limited information on the association of social support and mortality in HIV-infected individuals in sub-Saharan Africa. Therefore further research and data are required to understand better the association between social support and mortality in HIV-infected individuals.
2. Findings from this study have shown further insight may be achieved through using a mixed methods analysis in assessing mortality risk factors in which modelling accounts

for unobserved variables that may be explained by qualitative studies. Most explanations in existing literature that have applied frailty modelling provide speculative reasons for the frailty variance. Hence more studies on risk factors for mortality in HIV-infected people using mixed methods are required.

3. Using data from the Adult Wellness study, risk factors for mortality were assessed per site and the findings compared. Participants enrolled at each site were from different townships (localities) in the neighbourhoods surrounding the study sites. Using the participant's recorded street address at enrolment, it would be possible to determine their geo-spatial location. This would allow for the assessment of risk factors for mortality while accounting for unobserved variables and spatial effects. Geo-spatial analysis could be helpful in identifying hot spots and where clusters of risk factors for mortality exist.
4. Based on findings from the qualitative study, future quantitative studies looking at mortality in HIV infected people should include questions or tools relating to social support that have been validated in sub-Saharan Africa.

References

1. Hanagal DD, Dabade AD. A comparative study of shared frailty models for kidney infection data with generalized exponential baseline distribution. Journal of Data Science. 2013;11:109-42.

2. Hirsch K, Wienke A. Software for semiparametric shared gamma and log-normal frailty models: An overview. Computer Methods And Programs In Biomedicine. 2012;107:582-97.

3. Hougaard P. Frailty models for survival data. Lifetime Data Anal. 1995;1(3):255-73.

4. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography. 1979;16(3):439-54.

5. Wienke A. Frailty models in survival analysis. USA: Chapman and Hall/CRC; 2011. 301 p.

6. Dias SS, Andreozzi V, Martins MO, Torgal J. Predictors of mortality in HIV-associated hospitalizations in Portugal: a hierarchical survival model. BMC Health Serv Res. 2009;9:125. Epub 2009/07/25.

7. Hanagal DD. Modeling survival data using frailty models. USA: Chapman and Hall/CRC; 2011.

8. Manda SOM. A Nonparametric Frailty Model for Clustered Survival Data. Commun Stat-Theor M. 2011;40(5):863-75.

9. Alemu AW, Sebastian MS. Determinants of survival in adult HIV patients on antiretroviral therapy in Oromiyaa, Ethiopia. Global Health Action. 2010;3(5398).

10. Erikstrup C, Kallestrup P, Zinyama R, Gomo E, Mudenge B, Gerstoft J, et al. Predictors of mortality in a cohort of HIV-1-infected adults in rural Africa. Journal of Acquired Immune Deficiency Syndromes. 2007;44(4):478-83. Epub 2007/01/30.

11. Poka-Mayap V, Pefura-Yone EW, Kengne AP, Kuaban C. Mortality and its determinants among patients infected with HIV-1 on antiretroviral therapy in a referral centre in Yaounde, Cameroon: a retrospective cohort study. BMJ open. 2013;3(7).

12. Tadesse K, Haile F, Hiruy N. Predictors of mortality among patients enrolled on antiretroviral therapy in Aksum Hospital, Northern Ethiopia: a retrospective cohort study. PLoS One. 2014;9(1):e87392.

13. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society (Series B). 1972;34(2):187-220.

14. UNAIDS. Global report: UNAIDS report on the global AIDS epidemic. Accessed on 27th February 2014 at http://www.unaids.org/sites/default/files/media_asset/UNAIDS_Global_Report_2013_en_1.pdf. 2013.

15. Hoffmann CJ, Fielding KL, Johnston V, Charalambous S, Innes C, Moore RD, et al. Changing predictors of mortality over time from cARTstart: Implications for care. Journal of Acquired Immune Deficiency Syndromes. 2011(58):269-76.

16. Russell EC, Charalambous S, Pemba L, Churchyard GJ, Grant AD, Fielding K. Low haemoglobin predicts early mortality among adults starting antiretroviral therapy in an HIV care programme in South Africa: a cohort study. BMC Public Health. 2010;10:433. Epub 2010/07/27.

17. UNAIDS. Global report: UNAIDS report on the global AIDS epidemic 2010. Accessed on 31st December 2015 at http://www.unaids.org/globalreport/documents/20101123_GlobalReport_full_en.pdf.

18. UNAIDS. World AIDS Day Report 2011. Accessed on 31st December 2015 at http://www.unaids.org/sites/default/files/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2216_WorldAIDSday_report_2011_en.pdf.

19. Newell ML, Coovadia H, Cortina-Borja M, Rollins N, Gaillard P, Dabis F. Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis. Lancet. 2004;364(9441):1236-43. Epub 2004/10/07.

20. WHO. Global HIV/AIDS Response. Accessed on 31st December 2015 at http://www.who.int/hiv/pub/progress_report2011/en/. 2011.

21. Johannessen A, Naman E, Ngowi BJ, Sandvik L, Matee MI, Aglen HE, et al. Predictors of mortality in HIV-infected patients starting antiretroviral therapy in a rural hospital in Tanzania. BMC Infect Dis. 2008;8:52. Epub 2008/04/24.

22. Rajagopalan N, Suchitra JB, Shet A, Khan ZK, Martin-Garcia J, Nonnemacher MR, et al. Mortality among HIV-infected patients in resource limited settings: A case controlled analysis of inpatients at a community care center. American Journal of Infectious Diseases. 2009;5(3):226-31.

23. Wamalwa DC, Obimbo EM, Farquhar C, Richardson BA, Mbori-Ngacha DA, Inwani I, et al. Predictors of mortality in HIV-1 infected children on antiretroviral therapy in Kenya: a prospective cohort. BMC Pediatr. 2010;10:33. Epub 2010/05/21.

24. Harrell FE. Regression modeling strategies. United States of America: Springer-Verlag; 2001.

25. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S. A comparison of goodness-of-fit tests for the logistic regression model. Statistics in medicine. 1997;16(9):965-80. Epub 1997/05/15.

26. Hosmer DW, Lemeshow S. Applied logistic regression: John Wiley and Sons, Inc.; 2000.

27. Klein JP, Moeschberger ML. Survival analysis techniques for censored and truncated data: Springer; 2005.

28. Bhaskar R. Philosophy and scientific realism. In: *Critical Realism: Essential Readings, Critical Realism: Interventions (eds M. Archer, R. Bhaskar, A. Collier, T.Lawson & A. Norrie)*. London: Routledge; 1998.

29. Bergin M, Wells JSG, Owen S. Critical realism: a philosophical framework for the study of gender and mental health. Nurs Philos. 2008;9(3):169-79.

30. Clark MA, MacIntyre PD, Cruickshank J. A critical realist approach to understanding and evaluating heart health programmes. Health. 2007;11(4):513-39.

31. Yeung HW. Critical realism and realist research in human geography: a method or a philosophy in search of a method? Progress in Human Geography. 1997;21(1):51-74.

32. Al-Benna S, Al-Ajam Y, Way B, Steinstraesser L. Descriptive and inferential statistical methods used in burns research. Burns : journal of the International Society for Burn Injuries. 2010;36(3):343-6.

33. Altman DG. Statistical reviewing for medical journals. Statistics in medicine. 1998;17:2661-74.

34. Reed III JF, Salen P, Bagher P. Methodological and Statistical Techniques: What do residents really need to know about statistics? . Journal of Medical Systems. 2003;27:233-8.

35. Strasak AM, Zaman Q, Marinell G, Pfeiffer PK, Ulmer H. The use of statistics in medical research: A comparison of the New England Journal of Medicine and Nature Medicine. The American Statistician 2007;61:47-55.

36. Taback N, Krzyzanowska MK. A survey of abstracts of high-impact clinical journals indicated most statistical methods presented are summary statistics. Journal of Clinical Epidemiology. 2008;61:277-81.

37. Agaba PA, Digin E, Makai R, Apena L, Agbaji OO, Idoko JA, et al. Clinical characteristics and predictors of mortality in hospitalized HIV-infected Nigerians. Journal of infection in developing countries. 2011;5(5):377-82.

38. Ledwaba L, Tavel JA, Khabo P, Maja P, Qin J, Sangweni P, et al. Pre-ART levels of inflammation and coagulation markers are strong predictors of death in a South African cohort with advanced HIV disease. PLoS One. 2012;7(3):e24243.

39. Masiira B, Baisley K, Mayanja BN, Kazooba P, Maher D, Kaleebu P. Mortality and its predictors among antiretroviral therapy naive HIV-infected individuals with CD4 cell count >/=350 cells/mm(3) compared to the general population: data from a population-based prospective HIV cohort in Uganda. Glob Health Action. 2014;7:21843.

40. Altman DG, De Stavola BL, Love SB, Stepniewska KA. Review of survival analyses published in cancer journals. British Journal of Cancer. 1995;72(2):511-8.

41. Bakoyannis G, Touloumi G. Practical methods for competing risks data: a review. Statistical Methods in Medical Research. 2012;21(3):257-72.

42. Parpia S, Thabane L, Julian JA, Whelan TJ, Levine MN. Empirical comparison of methods for analyzing multiple time-to-event outcomes in a non-inferiority trial: a breast cancer study. BMC medical research methodology. 2013;13:44. Epub 2013/03/23.

43. Zhang X, Xu J, He J. Assessing non-inferiority with time-to-event data via the method of non-parametric covariance. Statistical Methods in Medical Research. 2011;22(3):346-60.

44. Duchateau L, Janssen P. Understanding heterogeneity in generalized mixed and frailty models. The American Statistician. 2005;59(2):143-6.

45. Rue H, Martino S, Chopin N. Approximate bayesian inference for latent gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society (Series B). 2009;71.

46. Martino S, Akerkar R, Rue H. Approximate Bayesian inference for survival models. Scandinavian Journal of Statistics. 2010;38:514-28.

47. Munda M, Rotolo F, Legrand C. parfm: Parametric Frailty Models in R. Journal of Statistical Software. 2012;51(11):1-20.

48. Duchateau L, Janssen P. Understanding heterogeneity in generalized mixed and frailty models. The American Statistician. 2005;59(2):143-6.

49. Lorino T, Sanaa M, Robin S, Daudin JJ. Comparison of semiparametric regression models for correlated survival data using simulations. Communication in Statistics-Theory and Methods. 2004;33(8):1975-91.

50. Manda SOM. A nonparametric frailty model for clustered survival data. Communications in Statistics - Theory and Methods. 2011;40(5):863-75.

51. Duchateau L, Janssen P. The Frailty Model. NY: Springer-Verlag; 2008.

52. Mietlowski W. A Review of: "Frailty Models in Survival Analysis, by A. Wienke". Journal of Biopharmaceutical Statistics. 2011;21(4):860-2.

53. Wienke A. Frailty models in survival analysis. USA: Chapman and Hall/CRC; 2011.

54. Otwombe KN, Petzold M, Martinson N, Chirwa T. A Review of the Study Designs and Statistical Methods Used in the Determination of Predictors of All-Cause Mortality in HIV-Infected Cohorts: 2002-2011. PLoS One. 2014;9(2):e87356. Epub 2014/02/06.

55. Martino S, Rue H. Implementing Approximate Bayesian Inference using Integrated Nested Laplace Approximation: a manual for the INLA program. Norway: Department of Mathematical Sciences, Norwegian University for Science and Technology, 2010.

56. Martins TG, Simpson D, Lindgren F, Rue H. Bayesian computing with INLA: new features. Computational Statistics and Data Analysis. 2013;67:68-83.

57. Ojikutu BO, Zheng H, Walensky RP, Lu Z, Losina E, Giddy J, et al. Predictors of mortality in patients initiating antiretroviral therapy in Durban, South Africa. South African Medical Journal. 2008;98(3):204-8. Epub 2008/03/20.

58. Otwombe KN, Petzold M, Modisenyane T, Martinson NA, Chirwa T. Factors associated with mortality in HIV-infected people in rural and urban South Africa. Glob Health Action. 2014;7:25488.

59. Helleberg M, Afzal S, Kronborg G, Larsen CS, Pedersen G, Pedersen C, et al. Mortality attributable to smoking among HIV-1-infected individuals: a nationwide, population-based cohort study. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2013;56(5):727-34.

60. City of Johannesburg. 2012/16 Integrated Development Plan. Accessed on 6th August 2014 at http://www.joburg-archive.co.za/2012/pdfs/idp/idp201216.pdf. 2012.

61. Shisana O, Simbayi LC, Zuma K, Jooste S, Zungu N, Labadarios D, et al. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Accessed on 18 April 2015 at http://www.hsrc.ac.za/uploads/pageContent/4565/SABSSM%20IV%20LEO%20final.pdf. 2014.

62. MacPherson P, Moshabela M, Martinson N, Pronyk P. Mortality and loss to follow-up among HAART initiators in rural South Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2009;103(6):588-93.

63. Martinson N, Variava E. Mortality in TB patients and active case finding in HAART treated adults. Accessed on 31 July 2014 at file:///C:/Users/otwombek/Downloads/TC_S___Dr_Neil_Martinson___Mortality_in_TB_patients_and_active_case_finding_in_HAART_treated_adults.pdf. 2008.

64. Wubshet M, Berhane Y, Worku A, Kebede Y, Diro E. High loss to followup and early mortality create substantial reduction in patient retention at antiretroviral treatment program in north-west ethiopia. ISRN AIDS. 2012;2012:721720.

65. Bushbuckridge Municipality. Bushbuckridge local municipality local development strategy. Accessed on 11 June 2014 at http://bushbuckridge.gov.za/wp-content/uploads/2014/03/Bushbuckridge-LED-Strategy-2010-20142.pdf. 2010.

66. Colditz GA, Emerson JD. The statistical content of published medical research: some implications for biomedical education. Medical Education. 1985;19(3):248-55.

67. Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. New England Journal of Medicine. 1983;309:709-13.

68. Kleinman K, Horton N. New stuff in SAS 9.3-frailty models. Accessed on 13th April 2015 at http://sas-and-r.blogspot.co.za/2011/09/example-97-new-stuff-in-sas-93-frailty.html 2015.

69. Duchateau L, Janssen P. The Frailty Model: Springer; 2007.

70. Berg A, Meyer R, Yu J. Deviance information criterion for comparing stochastic volatility models. Journal of Business and Economic Statistics. 2004;22(107-120).

71. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society. 2002;64:583-639.

72. StataCorp. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP. 2015.

73. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2015.

74. Monaco JV, Gorfine M, Hsu L. General Semiparametric Shared Frailty Model: Estimation and Simulation with frailtySurv. 2017. p. 1-44.

75. Duchateau L, Janssen P. Penalized partial likelihood for frailties and smoothing splines in time to first insemination models for dairy cows. Biometrics. 2004;60(3):608-14.

76. Therneau TM, Grambsch PM, Pankratz VS. Penalized survival models and frailty. Journal of Computational and Graphical Statistics. 2003;12(1):156-75.

77. Ha ID, Noh M, Lee Y. frailtyHL: A Package for Fitting Frailty Models with H-likelihood. The R Journal. 2012;4(2):28-37.

78. Ha ID, Lee Y, Song JK. Hierarchical likelihood approach for frailty models. Biometrika. 2001;88:233-43.

79. Ha ID, Lee Y. Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models. Biometrika. 2005;92:717-23.

80. Martino S, Akerkar R, Rue H. Approximate Bayesian Inference for Survival Models. Scandinavian Journal of Statistics. 2011;38:514-28.

81. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). J Roy Statist Soc Ser B. 2009;71:319-92.

82. Rondeau V, Commenges D, Joly P. Maximum penalized likelihood estimation in a gamma-frailty model. Lifetime Data Analysis. 2003;9(2):139–53.

83. Joly P, Commenges D, Helmer C, Letenneur L. A penalized likelihood approach for an illness-death model with interval-censored data: application to age-specific incidence of dementia. Biostatistics. 2002;3(3):433-43.

84. Hanrahan CF, Golub JE, Mohapi L, Tshabangu N, Modisenyane T, Chaisson RE, et al. Body mass index and risk of tuberculosis and death. AIDS. 2010;24(10):1501-8. Epub 2010/05/28.

85. Venkatesh KK, de Bruyn G, Lurie MN, Mohapi L, Pronyk P, Moshabela M, et al. Decreased sexual risk behavior in the era of HAART among HIV-infected urban and rural South Africans attending primary care clinics. AIDS. 2010;24(17):2687-96. Epub 2010/09/03.

86. Little RJA. Regression with Missing X's: A Review. Journal of the American Statistical Association. 1992;87(420):1227-37.

87. Rubin DB. Multiple imputation after 18+ years. Journal of the American Statistical Association. 1996;91(434):473-89.

88. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Statistics in medicine. 1991;10(4):585-98. Epub 1991/04/01.

89. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC medical research methodology. 2013;13.

90. Smith JM, Firth J. Qualitative data analysis: the framework approach. Nurse Researcher. 2011;18(2):52-62.

91. Loh J. Inquiry into issues of trustworthiness and quality in narrative studies: a perspective. The Qualitaive Report 2013;18(65):1-15.

92. Morrow SL. Quality and trust worthiness in qualitative research in counseling psychology. Journal of Counseling Psychology. 2005;52(2):250-60.

93. Carcary M. The research audit trail-enhancing trust worthiness in qualitative inquiry. The Electronic Journal of Business Research Methods. 2009;7(1):11-24.

94. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Education for Information. 2004;22:63-75.

95. Wu S, Jin Z, Wei X, Gao Q, Lu J, Ma X, et al. Misuse of statistical methods in 10 leading Chinese medical journals in 1998 and 2008. The Scientific World Journal. 2011;11:2106-14.

96. Vahanikkila H, Nieminen P, Miettunen J, Larmas M. Use of statistical methods in dental research: comparison of four dental journals during a 10-year period. Acta Odontol Scand. 2009;67(4):206-11.

97. Elster AD. Use of Statistical-Analysis in the AJR and Radiology - Frequency, Methods, and Subspecialty Differences. Am J Roentgenol. 1994;163(3):711-5.

98. Lim KJ, Yoon DY, Yun EJ, Seo YL, Baek S, Gu DH, et al. Characteristics and Trends of Radiology Research: A Survey of Original Articles Published in AJR and Radiology between 2001 and 2010. Radiology. 2012;264(3):796-802.

99. Okeh UM. Statistical problems in medical research. Afr J Biotechnol. 2008;7(25):4819-26.

100. Rigby AS, Armstrong GK, Campbell MJ, Summerton N. A survey of statistics in three UK general practice journal. BMC medical research methodology. 2004;4(1):28.

101. Shuai P, Zhou XH, Lao LX, Li XS. Issues of design and statistical analysis in controlled clinical acupuncture trials: an analysis of English-language reports from Western journals. Statistics in medicine. 2012;31(7):606-18.

102. Manda SOM, Meyer R. Bayesian inference for recurrent events data using time-dependent frailty. Statistics in medicine. 2005;24(1263-1274).

103. Pennell ML, Dunson DB. Bayesian semiparametric dynamic frailty models for multiple event time data. Biometrics. 2006;62(4):1044-52.

104. Bellan SE, Fiorella KJ, Melesse DY, Getz WM, Williams BG, Dushoff J. Extra-couple HIV transmission in sub-Saharan Africa: a mathematical modelling study of survey data. Lancet. 2013;381(9877):1561-9.

105. Exavery A, Kante AM, Tani K, Hingora A, Phillips JF. Sociodemographic drivers of multiple sexual partnerships among women in three rural districts of Tanzania. Hiv/Aids. 2015;7:105-13.

106. Scorgie F, Chersich MF, Ntaganira I, Gerbase A, Lule F, Lo YR. Socio-demographic characteristics and behavioral risk factors of female sex workers in sub-saharan Africa: a systematic review. AIDS and behavior. 2012;16(4):920-33.

107. Serwadda D, Wawer MJ, Musgrave SD, Sewankambo NK, Kaplan JE, Gray RH. HIV risk factors in three geographic strata of rural Rakai District, Uganda. AIDS. 1992;6(9):983-9.

108. Shisana O, Zungu-Dirwayi N, Toefy Y, Simbayi LC, Malik S, Zuma K. Marital status and risk of HIV infection in South Africa. South African Medical Journal. 2004;94(7):537-43.

109. WHO. The World Health Report 2002. Reducing Risks, Promoting Healthy Life. Accessed on 17 June 2015 at http://www.who.int/whr/2002/en/whr02_en.pdf. 2002.

110. Mujugira A, Wester CW, Kim S, Bussmann S, Gaolathe T. Patients with advanced HIV type 1 infection initiating antiretroviral therapy in Botswana: treatment response and mortality. AIDS Research and Human Retroviruses. 2009;25(2):127-33.

111. Otwombe KN, Laher F, Tutu-Gxashe T, Gray G, Mohapi L. The effect of a maturing antiretroviral program on early mortality for patients with advanced immune-suppression in Soweto, South Africa. PLoS One. 2013;8(11):e81538.

112. Ahsan Ullah AK. HIV/AIDS-Related Stigma and Discrimination: A Study of Health Care Providers in Bangladesh. Journal of the International Association of Physicians in AIDS Care. 2011;10(2):97-104.

113. Famoroti TO, Fernandes L, Chima SC. Stigmatization of people living with HIV/AIDS by healthcare workers at a tertiary hospital in KwaZulu-Natal, South Africa: a cross-sectional descriptive study. BMC medical ethics. 2013;14 Suppl 1:S6.

114. Horsman JM, Sheeran P. Health care workers and HIV/AIDS: a critical review of the literature. Social science & medicine. 1995;41(11):1535-67.

115. Knussen C, Niven CA. HIV AIDS and health care workers: Contact with patients and attitudes towards them. Psychol Health. 1999;14(3):367-78.

116. Holzemer WL, Makoae LN, Greeff M, Dlamini PS, Kohi TW, Chirwa ML, et al. Measuring HIV stigma for PLHAs and nurses over time in five African countries. SAHARA J : journal of Social Aspects of HIV/AIDS Research Alliance / SAHARA , Human Sciences Research Council. 2009;6(2):76-82.

117. Letamo G. The discriminatory attitudes of health workers against people living with HIV. PLoS medicine. 2005;2(8):e261.

118. Cotton S, Puchalski CM, Sherman SN, Mrus JM, Peterman AH, Feinberg J, et al. Spirituality and religion in patients with HIV/AIDS. Journal of general internal medicine. 2006;21 Suppl 5:S5-13.

119. Trevino KM, Pargament KI, Cotton S, Leonard AC, Hahn J, Caprini-Faigin CA, et al. Religious coping and physiological, psychological, social, and spiritual outcomes in patients with HIV/AIDS: cross-sectional and longitudinal findings. AIDS and behavior. 2010;14(2):379-89.

120. Belisle HA, Hennink M, Ordonez CE, John S, Ngubane-Joye E, Hampton J, et al. Concurrent use of traditional medicine and ART: Perspectives of patients, providers and traditional healers in Durban, South Africa. Glob Public Health. 2015;10(1):71-87.

121. Puoane TR, Hughes GD, Uwimana J, Johnson Q, Folk WR. Why HIV positive patients on antiretroviral treatment and/or cotrimoxazole prophylaxis use traditional medicine: perceptions of health workers, traditional healers and patients: a study in two provinces of South Africa. African journal of traditional, complementary, and alternative medicines : AJTCAM / African Networks on Ethnomedicines. 2012;9(4):495-502.

122. Babb DA, Pemba L, Seatlanyane P, Charalambous S, Churchyard GJ, Grant AD. Use of traditional medicine by HIV-infected individuals in South Africa in the era of antiretroviral therapy. Psychology, health & medicine. 2007;12(3):314-20.

123. Lemke S. Nutrition security, livelihoods and HIV/AIDS: implications for research among farm worker households in South Africa. Public health nutrition. 2005;8(7):844-52.

124. Twine W, Hunter LM. Adult mortality and household food security in rural South Africa: Does AIDS represent a unique mortality shock? Development Southern Africa. 2011;28(4):431-44.

125. Gregson S, Mushati P, Nyamukapa C. Adult mortality and erosion of household viability in AIDS-afflicted towns, estates, and villages in eastern Zimbabwe. J Acquir Immune Defic Syndr. 2007;44(2):188-95.

126. Sartorius B, Kahn K, Collinson MA, Sartorius K, Tollman SM. Dying in their prime: determinants and space-time risk of adult mortality in rural South Africa. Geospatial health. 2013;7(2):237-49.

127. Dowdy DW, Geng EH, Christopoulos KA, Kahn JS, Hare CB, Wlodarczyk D, et al. Mortality Among Antiretroviral-Eligible Patients in an Urban Public Clinic. Journal of Acquired Immune Deficiency Syndromes. 2011;57(4):297-300.

128. Kalichman SC, Cherry C, Amaral C, White D, Kalichman MO, Pope H, et al. Health and Treatment Implications of Food Insufficiency among People Living with HIV/AIDS, Atlanta, Georgia. J Urban Health. 2010;87(4):631-41.

129. Weiser SD, Fernandes KA, Brandson EK, Lima VD, Anema A, Bangsberg DR, et al. The Association Between Food Insecurity and Mortality Among HIV-Infected Individuals on HAART. Journal of Acquired Immune Deficiency Syndromes. 2009;52(3):342-9.

130. Kabore I, Bloem J, Etheredge G, Obiero W, Wanless S, Doykos P, et al. The effect of community-based support services on clinical efficacy and health-related quality of life in HIV/AIDS patients in resource-limited settings in sub-Saharan Africa. AIDS patient care and STDs. 2010;24(9):581-94.

131. Knight LC, Van Rooyen H, Humphries H, Barnabas RV, Celum C. Empowering patients to link to care and treatment: qualitative findings about the role of a home-based HIV counselling, testing and linkage intervention in South Africa. AIDS Care. 2015:1-6.

132. Massavon W, Barlow-Mosha L, Mugenyi L, McFarland W, Gray G, Lundin R, et al. Factors Determining Survival and Retention among HIV-Infected Children and Adolescents in a Community Home-Based Care and a Facility-Based Family-Centred Approach in Kampala, Uganda: A Cohort Study. ISRN AIDS. 2014;2014:852489.

133. Root R, Whiteside A. A qualitative study of community home-based care and antiretroviral adherence in Swaziland. Journal of the International AIDS Society. 2013;16:17978.

134. Young SD, Hlavka Z, Modiba P, Gray G, Van Rooyen H, Richter L, et al. HIV-related stigma, social norms, and HIV testing in Soweto and Vulindlela, South Africa: National Institutes of Mental Health Project Accept (HPTN 043). J Acquir Immune Defic Syndr. 2010;55(5):620-4.

135. Mohanan P, Kamath A. Family support for reducing morbidity and mortality in people with HIV/AIDS. The Cochrane database of systematic reviews. 2009(3):CD006046.

OPEN ACCESS Freely available online

PLOS ONE

A Review of the Study Designs and Statistical Methods Used in the Determination of Predictors of All-Cause Mortality in HIV-Infected Cohorts: 2002–2011

Kennedy N. Otwombe[1,2]*, Max Petzold[3], Neil Martinson[1,4], Tobias Chirwa[2]

1 Perinatal HIV Research Unit, University of the Witwatersrand, Johannesburg, South Africa, **2** School of Public Health, University of the Witwatersrand, Johannesburg, South Africa, **3** Centre for Applied Biostatistics, Occupational and Environmental Medicine, University of Gothenburg, Gothenburg, Sweden, **4** Johns Hopkins University, Center for TB Research, Baltimore, Maryland, United States of America

Abstract

Background: Research in the predictors of all-cause mortality in HIV-infected people has widely been reported in literature. Making an informed decision requires understanding the methods used.

Objectives: We present a review on study designs, statistical methods and their appropriateness in original articles reporting on predictors of all-cause mortality in HIV-infected people between January 2002 and December 2011. Statistical methods were compared between 2002–2006 and 2007–2011. Time-to-event analysis techniques were considered appropriate.

Data Sources: Pubmed/Medline.

Study Eligibility Criteria: Original English-language articles were abstracted. Letters to the editor, editorials, reviews, systematic reviews, meta-analysis, case reports and any other ineligible articles were excluded.

Results: A total of 189 studies were identified (n = 91 in 2002–2006 and n = 98 in 2007–2011) out of which 130 (69%) were prospective and 56 (30%) were retrospective. One hundred and eighty-two (96%) studies described their sample using descriptive statistics while 32 (17%) made comparisons using t-tests. Kaplan-Meier methods for time-to-event analysis were commonly used in the earlier period (n = 69, 76% vs. n = 53, 54%, p = 0.002). Predictors of mortality in the two periods were commonly determined using Cox regression analysis (n = 67, 75% vs. n = 63, 64%, p = 0.12). Only 7 (4%) used advanced survival analysis methods of Cox regression analysis with frailty in which 6 (3%) were used in the later period. Thirty-two (17%) used logistic regression while 8 (4%) used other methods. There were significantly more articles from the first period using appropriate methods compared to the second (n = 80, 88% vs. n = 69, 70%, p-value = 0.003).

Conclusion: Descriptive statistics and survival analysis techniques remain the most common methods of analysis in publications on predictors of all-cause mortality in HIV-infected cohorts while prospective research designs are favoured. Sophisticated techniques of time-dependent Cox regression and Cox regression with frailty are scarce. This motivates for more training in the use of advanced time-to-event methods.

Citation: Otwombe KN, Petzold M, Martinson N, Chirwa T (2014) A Review of the Study Designs and Statistical Methods Used in the Determination of Predictors of All-Cause Mortality in HIV-Infected Cohorts: 2002–2011. PLoS ONE 9(2): e87356. doi:10.1371/journal.pone.0087356

Editor: Barbara Ensoli, Istituto Superiore di Sanità, Italy

Received November 25, 2013; **Accepted** December 19, 2013; **Published** February 3, 2014

Funding: KO was supported by an academic scholarship from the Canadian-Africa Prevention Trials Network, the Consortium for Advanced Research Training in Africa (Carnegie Corporation of New York Grant no. B8606.R01, Swedish International Development Corporation Agency Grant no. 54100029, Ford Foundation Grant no. 1120-1838 and Wellcome Trust Grant no. 087547/Z/08/Z) and the Forgaty International Center (grant no. TW007370/3). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: otwombek@phru.co.za

Introduction

Appropriate utilization of biostatistical methods is becoming increasingly important in biomedical research. Many journals, if not all, have a dedicated statistical committee that scrutinizes the methods used in analyzing data. In the last decade, several papers addressing study design issues and statistical analysis approaches in different clinical fields have been published underpinning the importance of robustness in methodology [1–8]. There is consensus that inappropriate study designs and statistical methodology lead to incorrect results, poor interpretation of study findings and wrong conclusions.

An array of study designs and appropriate statistical techniques with varying levels of complexity exists. Selecting the appropriate study design and relevant statistical analysis technique is largely dependent on the complexity of the study and its objectives. Research on statistical content of medical research shows wider usage of techniques [9,10] beyond descriptive statistics as a result of advanced software that can handle complex analyses. Much as advanced analyses are being conducted, simple techniques of

Table 1. Classification of statistical methods as reported in journals.

Category	Brief description
***t*-tests**	One-sample, matched pair and two-sample *t*-tests
Contingency tables	Chi-square tests, Fisher's exact test, McNemar's test
Pearson correlation	Classical product moment-correlation
Simple linear regression	Least squares regression with one predictor and one response variable
Power	Loosely defined, includes use of the size of detectable (or useful) difference in determining sample size
Epidemiological statistics	Relative risk, odds ratio, log odds, measures of association, sensitivity, specificity
Adjustment and standardization	Pertains to incidence rates and prevalence rates
Multiway tables	Mantel-Haenszel procedure, log-linear models
Non-parametric tests	Sign test, Wilcoxon signed-rank test, Mann-Whitney test
Non-parametric correlation	Spearman's *rho*, Kendall's *tau*, test for trend
Analysis of variance	Analysis of variance, analysis of co-variance and *F*-tests
Multiple comparisons	Procedures for handling multiple inferences on same data sets, includes Bonferroni techniques, Scheffe's contrasts, Duncan multiple range procedures, Newmann-Keuls procedure
Transformation	Use of data transformation (e.g logarithms) often in regression
Multiple regression	Includes polynomial regression and stepwise
Life table	Actuarial life table, Kaplan-Meier estimate of survival
Regression for survival	Includes Cox regression and logistic regression
Other survival analysis	Breslow's Kruskal-Wallis, log-rank, Cox model for comparing survival
Cost benefit analysis	The process of combining estimates of cost and health outcomes to compare policy alternatives
Sensitivity analysis	Examines sensitivity of outcome to modest changes in parameters of model, or in other assumptions
Other	Anything not fitting above headings, includes cluster analysis, discriminant analysis and some mathematical modeling

doi:10.1371/journal.pone.0087356.t001

descriptive and inferential statistical analysis like student t-tests and chi-square tests remain popular in the literature [4,6,11].

Despite major successes in the development of interventions for prevention of mother to child treatment (PMTCT) and anti-retrovirals (ARVs), HIV still remains a major public health concern. To date, limited information is available if any, reporting on the study design and statistical techniques used in determining the predictors of all-cause mortality in HIV positive cohorts in the last decade. With a large number of clinicians and public health experts relying on published research for new developments in HIV research, it is important they understand appropriateness of study designs and statistical techniques used in determining predictors of all-cause mortality. This study reviews relevant original articles in HIV-infected cohorts with the aim of identifying study designs, statistical methods used and further assess their appropriateness. We also sought to determine whether there was an increase in the use of time-to-event analysis techniques over time and highlight the need for methodological training.

Methods

Search strategy and selection criteria

In this bibliometric analysis, we searched all original English-language articles indexed in Pubmed/Medline using the terms "Predictors of HIV Mortality", "Determinants of HIV Mortality" and "Factors associated with HIV mortality". The search covered the period between January 2002 and December 2011, a period of ten years. These were further split into two five year periods; January 2002–December 2006 and January 2007 to December 2011 in order to assess whether there was a variation in the methods used over time. Original articles on HIV-infected cohorts within the specified period were eligible for inclusion. Letters to the editor, editorials, reviews, systematic reviews, meta-analysis and case reports were excluded. Other studies comparing both HIV positive and negative participants were also excluded. We identified a total of 91 and 98 papers between the periods 2002–2006 and 2007–2011 respectively.

Each article was reviewed to determine the study design, nature of statistical methods used and their appropriateness. Time-to-event analysis methods were considered optimal or appropriate in this study. A spreadsheet containing a checklist of items of interest was prepared as a data collection tool. Findings were systematically recorded based on statistical methods previously reported [12]. We used a modified version of the classification proposed by Colditz and Emerson (Table 1) [13,14]. Where a statistical technique was used more than once in an article, we recorded it as having occurred only once. A count of the number of statistical techniques employed in each article was determined for purposes of comparing the two periods.

The statistical methods used in the research articles were classified as either parametric or non-parametric. A further classification was made describing the statistical methods used as either basic or advanced. Methods classified as basic included Student t-test, Chi-Square and Fishers Exacts test, Mann-Whitney, Kruskall-Wallis, Wilcoxon, simple one-way ANOVA and correlation statistics. Modelling approaches such as logistic Regression, Conditional Logistic Regression, Poisson Regression, Cox regression, time-varying Cox-regression and Cox regression with frailty and epidemiologic statistics were classified as advanced (Table 1).

The logistic regression is used to analyse the relationship between a binary dependent variable and independent predictor through estimation of the probability of an event occurring. It

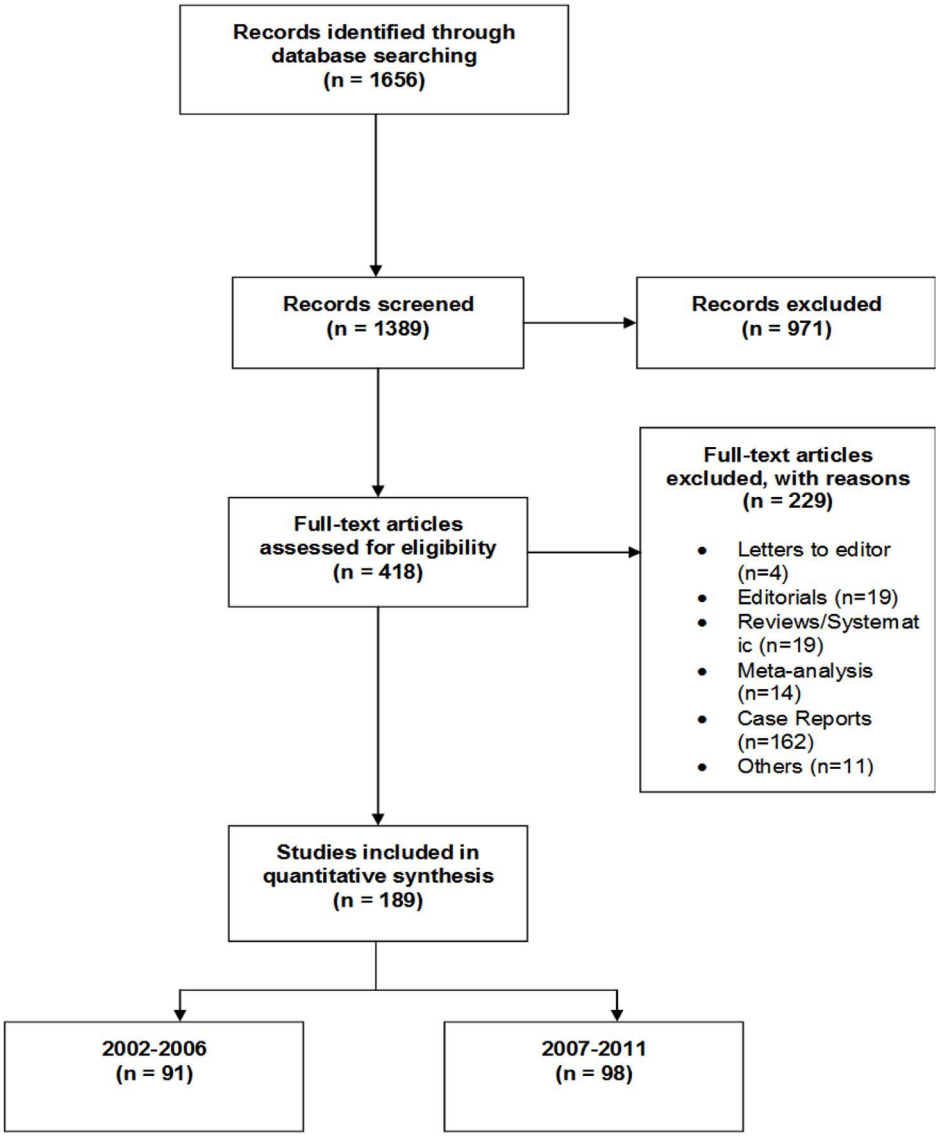


Figure 1. Flow chart showing the selection process of articles and the number in each period.
doi:10.1371/journal.pone.0087356.g001

makes no assumption about normality, linearity and homogeneity of variance. But used with time-to-event outcomes, it fails to account for follow-up time. For this reason, articles reporting use of logistic regression on such outcomes were classified as sub-optimal [15]. Cox regression analysis is a survival analysis technique in time-to-event data that incorporates follow-up time and fixed covariates [16]. Censoring is done when events occur. The method assumes risk of an event is homogeneous. Extensions of the Cox regression exist which include time dependent Cox regression and Cox regression analysis with frailty [17,18]. Time dependent Cox regression analysis accounts for the inherent correlation that may exist when covariates change over time. Cox regression analysis with frailty, if used in some of the reviewed articles, tries to account for unobserved heterogeneity.

The data collected in this study were compared between two periods. Frequency analysis was used to determine the number of studies reporting use of specified statistical techniques. The number of optimal or sub-optimal methods used in the determination of predictors of mortality was determined using frequencies. The comparison between numbers of methods reported between the two periods was compared using the chi-square test where appropriate. All the Statistical analysis was performed using SAS 9.3 software and p-values ≤ 0.05 were considered a significant difference.

Results

The total number of studies reporting on predictors of HIV mortality that met our criteria in the era January 2002 and December 2011 was 189 (n = 91 in 2002–2006 and n = 98 in 2007–2011). Figure 1 is a flow chart displaying the selection criteria that was followed in arriving at the final number of articles. All the identified articles used at least one (basic or advanced) statistical test. Journal of Acquired Immune Infection (JAIDS) (n = 34, 18%) and AIDS (n = 24, 13%) published more articles on HIV mortality. JAIDS and AIDS published 19/34 (56%) and 14/24 (58%) of these articles in the era 2002–2006. Majority of the articles used a prospective study design and the number was

Table 2. Types of study designs and sample sizes.

	2002–2006		2007–2011		
Research Design	**No.**	**(%)**	**No.**	**(%)**	**p-value**
Prospective	67	74	62	63	0.66
Retrospective	23	25	33	34	0.14
Case control	1	1	3	3	-
Prospective	1	-	0	-	-
Retrospective	0	-	3	-	-
Total	91		98		
Sample sizes					
up to 200	22	24	15	15	0.5
201 to 500	21	23	25	26	
501 to 1,000	16	18	20	20	
>1,000	32	35	38	39	

doi:10.1371/journal.pone.0087356.t002

similar in both periods (n = 67, 74% vs. n = 62, 63%; p = 0.66). Sample sizes varied from under 200 to greater than 1,000 participants. Table 2 presents the study design and sample size distribution of the included articles between the two periods. There were no significant differences in the study designs and sample sizes used between the two periods.

The number of studies reporting descriptive statistics for the two periods was similar. Table 3 presents the distribution of commonly reported statistical methods. The number of articles reporting use of t-tests, contingency table analysis, correlation and epidemiological statistics was similar. Few studies in both periods reported using one-way analysis of variance technique (ANOVA). The number of modeling approaches such as logistic, conditional logistic, generalized estimating equations and Poisson regression was significantly higher in the later period compared to the earlier (n = 31, 32% vs. n = 13, 14% p = 0.005).

A total of 122 (65%) articles reported using the Kaplan-Meier methods and it was commonly used in the first period compared to the second (p = 0.002). Use of the Cox proportional hazards regression modeling was reported by 131 (69%) articles and the number was similar between the two eras (p = 0.12). The number reporting use of time dependent Cox regression was higher in the first period (n = 21, 23% vs. n = 11, 11%; p = 0.03). Overall Cox regression with frailty was scarcely used (n = 7, 4%) in which 6 (3%) articles were in the later period.

There were 22 (12%), 96 (51%) and 71 (38%) articles reporting use of 2 to 3, 4 to 5 and more than 5 statistical methods respectively. There were no significant differences in the number of methods used between the two eras. Similarly there were no significant differences between the two eras in the number of

Table 3. Summary of statistical methods.

	2002–2006		2007–2011		
Statistical methods	**No. of publications (N = 91)**	**(%)**	**No. of publications (N = 98)**	**(%)**	**p-value**
Descriptive statistics	88	97	94	96	0.78
Inferential methods					
t-test	13	14	19	19	0.35
Contingency table analysis					
Basic (chi-square/Fisher exact)	25	27	37	38	0.13
Correlation	9	10	1	1	-
Epidemiologic statistics	39	43	45	46	0.67
Analysis of variance	3	3	1	1	-
Regression					
#Multiple regression	13	14	31	32	0.005
Survival analysis					
Kaplan-Meier	69	76	53	54	0.002
Cox-Regression	68	75	63	64	0.12
Time dependent Cox regression	21	23	11	11	0.03
Cox-Regression with frailty	1	1	6	6	-
Non-parametric methods	86	95	91	93	0.64
No. of different inferential methods					
2 or 3 methods	10	11	11	11	0.96
4 or 5 methods	44	48	43	44	0.54
More than 5 methods	37	41	44	45	0.56
¶Basic analysis	36	40	45	46	0.38
*Advanced analysis	90	99	97	99	0.96

#Refers to logistic, conditional logistic, generalized estimating equations and Poisson regression methods.
¶Refers to Student t-test, Chi-Square and Fishers Exact test, Mann-Whitney, Kruskall-Wallis, Wilcoxon, simple ANOVA and correlation.
*Refers to Logistic Regression, Conditional Logistic Regression, Poisson Regression and Survival Analysis and Epidemiologic statistics.
doi:10.1371/journal.pone.0087356.t003

Table 4. Appropriateness of statistical methods for predictors of mortality.

	2002–2006		2007–2011	
	Prospective (N = 68)	Retrospective (N = 23)	Prospective (N = 62)	Retrospective (N = 36)
Statistical methods	**No. (%)**	**No. (%)**	**No. (%)**	**No. (%)**
Appropriate statistical analysis				
Cox regression	53 (78)	15 (65)	48 (77)	15 (42)
Time dependent Cox regression	16 (24)	5 (22)	10 (16)	1 (3)
Cox regression analysis with frailty	1 (1)	0 (0)	3 (5)	3 (8)
Poisson regression	3 (4)	0 (0)	4 (6)	0 (0)
Sub-Optimal statistical analysis				
Logistic regression	4 (6)	4 (17)	7 (11)	17 (47)
Unclear	1 (1)	2 (9)	2 (3)	3 (8)

Note: Totals in this table do not add up to the number of articles because some articles used more than one method in their analysis.
doi:10.1371/journal.pone.0087356.t004

articles reporting use of basic or advanced statistical analysis methods.

A total of 149 (79%) of the articles used appropriate methods while 40 (21%) used sub-optimal methods to determine the predictors of mortality in HIV-infected participants. Of the articles using appropriate methods, 116 (78%) were prospective and 33 (22%) retrospective. There were significantly more articles from the first period using appropriate methods compared to the second (n = 80, 88% vs. n = 69, 70%, p-value = 0.003). Table 4 presents findings on the appropriateness of the statistical methods used. A significantly higher number of articles in the first period could have used Cox regression analysis with frailty as the appropriate method, since they had clustered data (n = 82, 92% vs. n = 65, 68%; $p<0.0001$) while overall they were 147 (78%).

Discussion

This paper aimed at reviewing articles on predictors of all-cause mortality in HIV-infected people to investigate the appropriateness of statistical methods used and nature of study designs. We reviewed a total of 189 articles. Like in any other study, there were several limitations. The literature review of the articles included in this study was searched in Pubmed/Medline ostensibly because this was not a systematic review requiring a measure of effect. Any relevant articles indexed elsewhere or in a language other than English were not considered.

Our findings concur with others reporting on study designs, statistical methods used and their appropriateness. Prospective study designs remain the most common type of design used in studies of predictors of HIV mortality in the last decade. Retrospective study designs formed about one third of all articles included in this study. It may be that retrospective study designs are used as a cost-effective way of saving on huge expenses required for running prospective studies as a way for stimulating academic research. However there was no significant difference in the type of study designs used between the two periods.

Basic statistical analysis procedures like t-tests, Chi-Square and Fishers Exact tests, Mann-Whitney, Kruskall-Wallis and Wilcoxon are commonly used. There was no difference in the number of articles reporting use of t-tests between the two periods. This is similar to previously reported studies that have shown the popularity of t-tests [4,19].

All the studies used at least two statistical tests. We contend that our inclusion criteria and the nature of studies included in this review all required using a type of statistical analysis to address the research question. Our findings concur with those reported earlier showing majority of articles apply more than one statistical test [5,7,20]. But this is contrary to the findings of a review on study designs and statistical methods in Chinese journals that found a low proportion of studies reporting use of multiple statistical tests [21].

Survival analysis approaches remain popular in the studies looking at predictors of mortality in HIV-infected people, especially the Cox proportional hazards regression modeling. Though fewer studies used extensions of the Cox proportional hazards regression, our findings show that there is an interest in using advanced approaches like the time-dependent Cox proportional hazards or Cox proportional hazards regression with frailty in modeling survival data in HIV-infected cohorts. We found a higher proportion of the studies could have used Cox regression analysis with frailty, an appropriate technique. While the methods used were not wrong, they could have gained more information by using Cox regression analysis with frailty. Previously reported work on statistical methods in medical research show that while use of sophisticated methods is increasing, inappropriate techniques still remain a challenge [1,6,7,22]. It may be that recent techniques are advanced and require rigour to implement. Furthermore the techniques may not necessarily be easily implemented in standard statistical software [23]. As a result, researchers use techniques that are fairly straight-forward and implementable in standard statistical software.

Our findings show that not all the studies in our sample used optimal statistical tests in the determination of the predictors of mortality. Survival analysis techniques produce better estimates that are more informative when analysed using optimal methods. Furthermore, in clinical research where objectives require a multivariable analysis approach, it is prudent to adjust for confounding appropriately by using optimal statistical methods [24]. Cox regression analysis and its extensions provide a better picture compared to logistic regression when using survival data. Unlike previously reported research [21,25], the proportion of studies using sub-optimal statistical tests was lower in our sample. These findings are contrary to those reported in other clinical fields where there was a high proportion of articles using sub-optimal methods [3,11,19,26,27].

Descriptive statistics and survival analysis techniques remain the most common methods of analysis in publications on predictors of

all-cause mortality in HIV-infected cohorts while prospective research designs are favoured. These results suggest the importance of understanding advanced survival analysis methods in interpreting research findings in this set-up. However complex and appropriate methods like Cox regression analysis with frailty remain scarcely utilised. Our findings are in agreement with others who also reported a high use of descriptive statistics [4,6]. The more sophisticated techniques of time dependent Cox regression and Cox regression with frailty are scarcely used. This motivates for more training in the use of advanced time-to-event methods.

Supporting Information

Appendix S1 Prisma checklist.
(DOC)

Author Contributions

Conceived and designed the experiments: KO NM TC. Performed the experiments: KO. Analyzed the data: KO MP TC. Contributed reagents/materials/analysis tools: KO NM. Wrote the paper: KO MP NM TC.

References

1. Al-Benna S, Al-Ajam Y, Way B, Steinstraesser L (2010) Descriptive and inferential statistical methods used in burns research. Burns 36: 343–346.
2. Harris AHS, Reeder R, Hyun JK (2009) Common statistical and research design problems in manuscripts submitted to high-impact psychiatry journals: What editors and reviewers want authors to know. Journal of Psychiatric Research 43: 1231–1234.
3. Okeh UM (2008) Statistical problems in medical research. Africa Journal of Biotechnology 7: 4819–4826.
4. Reed III JF, Salen P, Bagher P (2003) Methodological and Statistical Techniques: What do residents really need to know about statistics? Journal of Medical Systems 27: 233–238.
5. Strasak AM, Zaman Q, Marinell G, Pfeiffer PK, Ulmer H (2007) The use of statistics in medical research: A comparison of the New England Journal of Medicine and Nature Medicine. The American Statistician 61: 47–55.
6. Taback N, Krzyzanowska MK (2008) A survey of abstracts of high-impact clinical journals indicated most statistical methods presented are summary statistics. Journal of Clinical Epidemiology 61: 277–281.
7. Wu S, Jin Z, Wei X, Gao Q, Lu J, et al. (2011) Misuse of statistical methods in 10 leading Chinese medical journals in 1998 and 2008. The Scientific World Journal 11: 2106–2114.
8. Strasak AM, Zaman Q, Pfeiffer PK, Gobel G, Ulma H (2007) Statistical errors in medical research-a review of common pitfalls. Swiss Medical Weekly 137: 44–49.
9. Altman DG (1998) Statistical reviewing for medical journals. Statistics in Medicine 17: 2661–2674.
10. Goldin J, Zhu W, Sayre JW (1996) A review of the statistical analysis used in papers published in clinical radiology. Clinical Radiology 51: 47–50.
11. Rigby AS, Armstrong GK, Campbell MJ, Summerton N (2004) A survey of statistics in three UK general practice journal. BMC Medical Research Methodology 4: 1–7.
12. Goldin J, Zhu W, Sayre JW (1996) A review of the statistical analysis used in papers published in clinical radiology. Clinical Radiology 51: 47–50.
13. Colditz GA, Emerson JD (1985) The statistical content of published medical research: some implications for biomedical education. Medical Education 19: 248–255.
14. Emerson JD, Colditz GA (1983) Use of statistical analysis in the *New England Journal of Medicine*. New England Journal of Medicine 309: 709–713.
15. Hosmer DW, Lemeshow S (2000) Applied logistic regression: John Wiley and Sons.
16. Cox DR (1972) Regression models and life-tables. Journal of the Royal Statistical Society Series B 34: 187–220.
17. Lin DY (1994) Cox regression analysis of multivariate failure time data: the marginal approach. Statistics in Medicine 13: 2233–2247.
18. Vaupel JW, Manton KG, Stallard E (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16: 439–454.
19. Elster AD (1994) Use of statistical analysis in the *AJR* and *Radiology*: Frequency, Methods and Subspeciality Differences. AJR 163: 711–715.
20. Jaykaran PY (2011) Quality of reporting statistics in two Indian pharmacology journals. Journal of Pharmacology and Pharmacotherapeutics 2: 85–90.
21. Wang Q, Zhang B (1998) Research design and statistical methods in Chinese medical journals. Journal of the American Medical Association 280: 283–285.
22. Anthony D (1996) A review of statistical methods in the *Journal of Advanced Nursing*. Journal of Advanced Nursing 24: 1089–1094.
23. Nietert PJ, Wahlquist AE, Herbert TL (2013) Characteristics of recent biostatistical methods adopted by researchers publishing in general/internal medicine journals. Statistics in Medicine 32: 1–10.
24. Vähänikkilä H, Niemine P, Miettunen J, Larmus M (2009) Use of statistical methods in dental research: comparison of four dental journals during a 10-year period. Acta Odontologica Scandinavica 67: 206–211.
25. Jin Z, Yu D, Zhang L, Meng H, Lu J, et al. (2010) A retrospective survey of research design and statistical analyses in selected Chinese medical journals in 1998 and 2008. Plos One 5: 1–4.
26. Lim KJ, Yoon DY, Yun EJ, Seo YL, Baek S, et al. (2012) A survey of original articles published in AJR and Radiology between 2001 and 2010. Radiology 264: 796–802.
27. Shuai P, Zhou X-H, Lao L, Li X (2012) Issues of design and statistical analysis in controlled clinical acupuncture trials: An analysis of English-language reports from Western Journals. Statistics in Medicine 31: 606–618.

Modeling clustered semiparametric frailty data: a comparison of methods using simulations

Kennedy N. Otwombe[1,2], Max Petzold[2,3], Neil Martinson[1,4], Tobias Chirwa[2]

Affiliations

[1]Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[3]Health Metrics, University of Gothenburg, Gothenburg, Sweden

[4]Johns Hopkins University, Center for TB Research, Baltimore, MD, USA

Author's Email Addresses

Kennedy N. Otwombe: otwombek@phru.co.za

Max Petzold: max.petzold@gu.se

Neil Martinson: martinson@phru.co.za

Tobias Chirwa: tobias.chirwa@wits.ac.za

Corresponding author

Kennedy Otwombe

Perinatal HIV Research Unit

Faculty of Health Sciences

University of the Witwatersrand

Johannesburg

South Africa

Email: otwombek@phru.co.za

Tel: +27 11 989 9807

Abstract

Background

Shared survival frailty models are commonly used to assess clustered time-to-event data where individuals enrolled into a study are followed until an event occurs or they are censored. While parametric frailty models are commonly used, semiparametric frailty models are increasingly being adopted. We compare four selected semiparametric frailty methods using different packages.

Methods

Using simulated clustered data, we compared the integrated nested Laplace approximation (INLA), h-likelihood (frailtyHL), penalized likelihood estimation (frailtyPack) and penalised partial likelihood (coxph). Performance of the methods is estimated using the mean squared error (MSE).

Results

In the simulations, INLA had a slightly better fit with consistently lower MSE values compared to the other methods in both gamma and log-normal frailty models. The random effects estimate for INLA, whose method is slightly different, was lower than other methods.

Conclusions

Although the MSE for INLA was slightly lower, the fixed-effects estimates were very similar with slight differences. All the methods handled the semiparametric frailty modelling well while others used slightly more computer resources with regards to time.

Keywords: Time-to-event; Heterogeneity; Semi-parametric Frailty;

Introduction

Time-to-event statistical methods are commonly used where time is an important consideration in relation to a primary outcome of interest. Here, Cox proportional hazards model is the preferred method in testing the effect of covariates on the hazard rates. It is defined as $h(t) = h_0(t)\exp(\beta X)$ where $h_0(t)$ is the baseline hazard, β is a vector of coefficients and X comprises the covariates [1]. It assumes independence between observations and linearity of covariates and the hazard ratio. It has the advantage of not assuming a distribution for $h_0(t)$ and being available in standard statistical software whereas its disadvantage includes not accounting for unobserved heterogeneity due to clustering.

Shared frailty models, the main focus of this study, are an extension of the Cox proportional hazards regression that account for unobserved heterogeneity in the presence of clustering [2, 3]. Examples include people living in one household, repeatedly collected data on an individual, a clinic or a school. These models are also suitable in situations where subjects in a study are more prone to failure than others. Initially introduced in lifetime models, extensions exist in several different fields. In time-to-event analysis, failure to account for unobserved heterogeneity may lead to biased estimates and incorrect interpretation of results. In frailty modelling, a random parameter is introduced in the hazard rate model to account for unobserved heterogeneity.

Frailty modelling assumes some form of distribution for the frailties. Several distributions for the frailties have been proposed in the literature with the most common being the gamma and log-normal. The gamma distribution is commonly used due to its mathematical tractability while the log-normal fits in well with the generalised linear mixed models [4]. Frailty models may be implemented in

parametric or semi-parametric forms. In the parametric form, a distribution is assumed for the survival time while in the non-parametric form, no distribution is assumed allowing for flexibility in the modelling process. While methods for modelling the parametric form are well developed, the semi-parametric form requires use of more complicated approaches [4, 5].

The comparison of various parametric shared frailty models using simulations has previously been done. The gamma and lognormal frailty distributions have been used to assess the correlation between variance estimates and the correlation of frailties [6]. Parametric frailty models fitted by maximum marginal likelihood with several possible baseline hazards have been illustrated to show how well they perform [4, 7]. A time-dependent frailty model within a Bayesian framework has previously been examined and compared with a constant frailty model using a recurrent events data [8]. Others have demonstrated the use of penalized likelihood to fit efficient shared frailty models [9]. Several frailty distributions, their statistical properties and various maximisation techniques are summarised elsewhere in excellent reviews on frailty models [10-13].

Semi-parametric frailty models have been used to compare various frailty distributions with the most common frailty distributions being the gamma and log-normal. Using a comparison of four models involving correlated data, it was shown that the semi-parametric frailty models perform better than the naïve Cox model [14]. The use of the Dirichlet process prior in the Bayesian modelling of frailties has been shown to perform better in the semi-parametric frailty modelling compared with the parametric [15]. Using simulations, several softwares capable of fitting shared semi-parametric frailty models were compared and found to fit well but with varying levels of complexity [16]. Other examples on the performance of various semiparametric

methods have been presented in literature using both gamma and lognormal frailties [10, 12, 13].

While several excellent examples on the use of semiparametric frailty models exist, research in the field continuous with more approaches being proposed. Additionally, several articles comparing how well semiparametric frailty methods perform exist. However, articles presenting an overview of maximisation methods in semiparametric frailty models that include the integrated nested Laplace approximation are limited. This study provides such an overview using the gamma and lognormal frailties through simulations.

Methods

Notation

Let $j = 1, \dots, n_i$ be subjects that belong to cluster i, where $i = 1, \dots, m$ and the endpoint of interest is a time-to-event outcome, for example death. Further let the observation time be T_{ij} where an event may or may not occur. In this case, the observation time may be defined as $T_{ij} = \min(T, C)$ where $C = 1$ if no censoring occurs and $C = 0$ when censored. The shared frailty model is defined as $h_{ij}(t) = h_0(t)\, exp\left(X_{ij}^t \beta + w_i\right)$ where $h_{ij}(t)$ is conditional hazard for subject j in cluster i, w_i is random effect defined as the frailty term, β is a vector of fixed effects or parameter estimates and X_{ij}^t is a vector of covariates from the observed data [10, 13]. This model is referred to as the shared frailty model because subjects in the same cluster are considered to share the same frailty. In the parametric form, a distribution is assumed for the time of events hence the baseline hazard $h_0(t)$ while in the semiparametric form, no distribution is specified.

In the semiparametric frailty model with gamma frailty, the frailty term is assumed to follow the gamma distribution. Here, we confine our study to the one-parameter gamma distribution $f_W(w) = \{w^{1/\theta-1}exp(-w/\theta)\}/\theta^{1/\theta}\Gamma(1/\theta)$ where the heterogeneity between clusters is assessed by θ [10]. Frailty values > 1 signify higher risk while < 1 suggests low risk. The gamma distribution is commonly used due to its ease in mathematical manipulations [10, 13].

The lognormal distribution is used due to its flexibility in modelling both random and fixed effects. It is defined as $f_W(w) = (1/w\sqrt{2\pi\gamma})\exp(-(\log w)^2/2\gamma)$ where γ is the variance and is always positive. While not as commonly used as the gamma distribution, the lognormal distribution is less tractable [10, 13]. The lognormal frailty model is restrictive in the quantification of the frailty variance due to its form [13]. Both gamma and lognormal frailty models analyse data using varying maximisation procedures. Below we briefly describe the maximisation procedures used in the methods adopted in this study.

Penalized partial likelihood (CoxPH)

In this approach, smoothing splines that help in parameter estimation through transformations, are used in conjunction with the penalized partial likelihood to determine parameter estimates. The penalized partial likelihood is given by $l^s_{PPL}(\beta) = log\left[\prod_{i=1}^{n}\prod_{j=1}^{n_i}\left[\exp\big(g(.)\big)/\sum_{R(t_{ij})}\exp(g(..))\right]^{\sigma_{ij}}\right] - \lambda\int\{g''(z)\}^2dz$ where the parameter estimate β defines functions $g(.)$, $g(..)$ and $g''(z)$ and $\lambda\int\{g''(z)\}^2dz$ is the penalty term subtracted from each iteration during the parametrization process [17, 18].

H-Likelihood (frailtyHL)

The h-likelihood functions fit both the gamma and lognormal frailty distributions. The approach utilises the Laplace approximation procedures leading to efficient parametrisation of frailty estimates. The approach is defined as $h = h(\beta, h_0, \alpha) = l_0 + l_1$ where l_0 and l_1 are conditional log densities and sum of log densities respectively [19]. An indepth analysis of the h-likelihood theory and its potential applications are presented elsewhere [19-21].

Piecewise log-constant baseline hazard (INLA)

In this Bayesian approach, the time of observation is partitioned and the baseline hazard in each partition is assumed to remain constant. Additionally, the effect of the covariates is assumed to remain constant. Here, parameter estimation is carried out using data augmentation through latent Gaussian models and through a first-order random walk (RW1) [22, 23]. The likelihood of the j^{th} observation in the i^{th} cluster is given by $\pi(y_{ij}|\eta_{ij}, \theta)$ where η_{ij} is a linear predictor and θ is a vector of hyper-parameters [23].

Penalized full Likelihood (frailtyPack)

This method uses a full likelihood in conjunction with splines in parameter estimation. Splines are useful where precise estimations in semiparametric frailty modelling are unachievable [24]. A penalty is imposed on the likelihood by this method through introduction of large values on functions classified as rough. It is defined as $pl(h_0(.), \beta, \theta) = l(h_0(.), \beta, \theta) - \sum_{d=1}^{K} K_d \int_0^{\infty} h''^2_{0d}(t)\, dt$ and $l(h_0(.), \beta, \theta)$ is the full likelihood and $K_d \geq 0$ are smoothing parameters per stratum [24, 25].

Simulation

We simulated the semiparametric frailty model with gamma and lognormal frailties following the frailtysurv R package approach [26] under the assumption that the gamma distribution has an $E(X) = 1$ and $Var(X) = \theta$ whereas the lognormal has an $E(\theta/2) = 1$ and $Var(X) = E(2\theta) - E(\theta)$. For each of the frailty models, 10 clusters of size 50 and 100 were generated 1,000 times. For each j^{th} observation that belongs to the i^{th} cluster, the conditional survival function was derived by $S_{ij}(t|Z_{ij}, w_i) = exp(-H_0(t)w_i exp(\beta X_{ij}))$ where the unspecified cumulative hazard function is represented by $H_0(t)$ and w_i and βX_{ij} are as previously defined [26]. To simulate the failure time T_{ij} in the region $(0, \infty)$, $H_0(t)$ is inverted and the failure times are determined by $T_{ij} = H_0^{-1}(-ln(U_{ij})\exp(-\beta X_{ij})/w_i)$ where U_{ij} is uniformly distributed. The generated frailties are dependent on the selected frailty distributions, in our case the gamma and lognormal. Using a normal distribution for right censoring, an indicator for failure time was defined as $\sigma_{ij} = I(T_{ij} \leq C_{ij})$ where C_{ij} is the censoring variable for the j^{th} subject in the i^{th} cluster. For purposes of illustration, we only present simulated data generated using 10 clusters that have a fixed censoring rate of 20%.

Statistical analysis

Simulated data were analysed by fitting semiparametric gamma and lognormal frailty models where parameter estimates and their standard errors were outputted. Additionally, frailty variances were generated for each model. The mean squared error (MSE) defined as $MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$ was used to compare parameter estimates across all simulated models. For comparison, box plots were generated for parameter estimates and standard errors for the models.

The semiparametric frailty models were fitted using the following R packages: Integrated Nested Laplace Approximation (*INLA*) for Piecewise log-constant baseline hazard, *frailtyHL* for H-likelihood, *coxPH* for penalized partial likelihood and *frailtyPack* for penalized full likelihood. The SAS SGPLOT procedure in SAS Enterprise Guide 7.1 (SAS Institute, Cary, NC) was used to generate plots.

Results

In the semiparametric gamma frailty models of cluster size 50 and 100 (figure 1a and 1b), the four methods with differing maximisation procedures produced similar results. Even though the parameter estimates were similar, INLA parameter estimates were slightly lower relative to those produced by other methods.

The lognormal frailty parameter estimates were also similar across all the methods for both clusters. In general, there was less variability in the clusters with more members (cluster size n=100) in comparison to that of cluster size n=50 (figures 1c and 1d).

The standard errors and INLA standard deviations are presented graphically in figure 2. In figure 2a, while the gamma plots appear similar, there seems to be greater variability with cluster size 50 compared to cluster size 100. In the lognormal plots, the standard errors/INLA standard deviation is similar across the methods. However the penalized full likelihood appears to have more variability under the lognormal frailty model with cluster size 100.

When comparing the MSE for each of the parameter estimates across all the four methods, they were all similar to 2 decimal points (table 1). Though a small difference, the MSE of INLA was consistently lower than other methods. Additionally, the MSE was lower with larger cluster sizes.

In the gamma frailty model considering both cluster sizes, the frailty variances were similar for penalized partial likelihood, H-likelihood and penalized full likelihood and ranged between 1.8 and 1.9 (table 2). The piecewise log-constant baseline hazard frailty variance was less. In the lognormal frailty, the frailty variances for the piecewise log-constant baseline hazard and penalized partial likelihood were similar and ranged between 1.1 and 1.2. Frailty variances estimated using the H-likelihood and penalized full likelihood methods were larger and ranged between 1.7 and 1.9.

Discussion

Our study has compared the performance of different maximisation procedures using simulations where parameter estimates and their standard errors/standard deviation were generated. Estimates from the four semiparametric frailty models including the Bayesian INLA were assessed by the MSE. While the estimates were generally similar, the MSE values showed that INLA had consistently lower values in comparison to the other methods. All the approaches were implemented with ease though some may take longer to iterate depending on the cluster size. This observation is similar to a previous study where several semiparametric frailty models were compared and found to be easily implementable but with varying computational complexities [16].

The semiparametric gamma frailty models are generally the most preferred due to their ease in implementation and appealing mathematical properties. In our study, the parameter estimates and standard errors/INLA standard deviation were similar. Additionally, less variability was observed with large clusters. Our findings are in harmony with other studies involving simulated gamma frailty models where the semiparametric gamma frailty model performed well [10, 13, 24]. Their standard errors were also similar across the methods. In the semiparametric lognormal frailty

modeling, parameter estimates were also similar. While the standard errors/INLA standard deviations were similar, some variability was observed in the penalized full likelihood approach.

The four methods compared in this study were all capable of analysing clustered semiparametric frailty models. Among them, the penalized partial likelihood was the fastest in iterating. This finding is in harmony with a previous study that compared several semiparametric frailty models [16]. Additionally, the four methods were compared using the MSE. Though the MSE findings were similar, the piecewise log-constant baseline hazard approach used by INLA produced lower values consistently.

The frailty variances were similar in a third of the methods assessed under the gamma frailty model and a half in the lognormal frailty. These differences are not unusual because the frailty variances in these approaches are estimated using different methods. Unfortunately only one approach using the H-Likelihood provides frailty variances that are accompanied with standard errors.

Limitations of this study include the complexity of accounting for unobserved heterogeneity which is conditional on latent variables. It is also important to correctly specify the frailty model. In addition, this study only focussed on four semiparametric methods that are available in the R packages. While other semiparametric frailty models exist such as those found in the R packages coxme and phmm amongst others, a detailed overview of their performance has been previously presented [10, 16, 19, 24, 26]. Using several different cluster sizes and censoring rates would have provided more diversity in the performance of these methods. However, and unlike previous research, this study adds INLA whose parameter estimation follows the

Bayesian theory (INLA) and is compared with the frequentist approaches using the MSE.

In conclusion, semiparametric frailty models can be implemented using any of these methods. However, one has to decide on an approach that utilises less computer time resources while at the same time providing optimal results. Our MSE findings suggest that INLA may be a worthy consideration if one is considering fitting a Bayesian model. Alternatively, the penalized partial likelihood approach would suffice for semiparametric frailty modelling using a frequentist approach.

Declaration section

List of abbreviations

Integrated nested Laplace approximation (INLA); Mean squared error (MSE);

Ethics and consent statement

Approval to proceed with this study was provided by the University of the Witwatersrand Human Research Ethics Committee.

Competing interests

The authors declare no competing interests.

Author's contributions

KO, NM, MP and TC conceived the study. KO performed literature review, statistical analysis and wrote the initial draft of the manuscript under the guidance of NM, MP and TC.

Availability of Data and Materials

The dataset supporting the conclusions of this article are available from the corresponding author on request.

Acknowledgement

We are grateful to the Perinatal HIV Research Unit (PHRU) for providing the Adult Wellness Study Data.

Funding

The Adult Wellness Study was funded by PEPFAR through USAID South Africa (grant no: 674-A-00-08-00009-00).

Analysis was supported by the Consortium for Advanced Research Training in Africa (CARTA). CARTA is jointly led by the African Population and Health Research Center and the University of the Witwatersrand and funded by the Wellcome Trust (UK) (Grant No: 087547/Z/08/Z), the Department for International Development (DfID) under the Development Partnerships in Higher Education (DelPHE), the Carnegie Corporation of New York (Grant No: B 8606), the Ford Foundation (Grant No: 1100-0399), Google.Org (Grant No: 191994), Sida (Grant No: 54100029) and MacArthur Foundation Grant No: 10-95915-000-INP.

KO received an academic scholarship from Canadian-Africa Prevention Trials Network and the Forgaty International Center (grant no. TW007370/3).

All the funding sources for this study had no role in the design, methods, analysis or manuscript preparation

References

1. Cox D: **Regression models and life-tables**. *Journal of the Royal Statistical Society (Series B)* 1972(43):187-220.
2. Clayton DG: **A model for association in bivariate life tables and its applications to epidemiological studies of familial tendency in chronic disease incidence**. *Biometrica* 1978, **65**:141-151.
3. Vaupel JW, Manton KG, Stallard E: **The impact of heterogeneity in individual frailty on the dynamics of mortality**. *Demography* 1979, **16**(3):439-454.
4. Munda M, Rotolo F, Legrand C: **parfm: Parametric Frailty Models in R**. *Journal of Statistical Software* 2012, **51**(11):1-20.
5. Duchateau L, Janssen P: **Understanding heterogeneity in generalized mixed and frailty models**. *The American Statistician* 2005, **59**(2):143-146.
6. Wienke A, Arbeev K, Locatelli I, Yashin AI: **A simulation study of different correlated frailty models and estimation strategies**. In.; 2003.
7. Rotolo F, Munda M, Callegaro A: **parfm: Parametric Frailty Models. R package version 2.7.1, URL** http://CRAN.R-project.org/package=parfm**.** In.; 2012.
8. Manda SOM, Meyer R: **Bayesian inference for recurrent events data using time-dependent frailty**. *Statistics in Medicine* 2005, **24**:1263-1274.
9. Therneau TM, Grambsch PM, Pankratz VS: **Penalized Survival Models and Frailty**. *Journal of Computational and Graphical Statistics* 2003, **12**(1):156-175.
10. Duchateau L, Janssen P: **The Frailty Model**. NY: Springer-Verlag; 2008.

11. Hanagal DD: **Modeling survival data using frailty models**. USA: Chapman and Hall/CRC; 2011.

12. Mietlowski W: **A Review of: "Frailty Models in Survival Analysis, by A. Wienke"**. *Journal of Biopharmaceutical Statistics* 2011, **21**(4):860-862.

13. Wienke A: **Frailty models in survival analysis**. USA: Chapman and Hall/CRC; 2011.

14. Lorino T, Sanaa M, Robin S, Daudin JJ: **Comparison of semiparametric regression models for correlated survival data using simulations**. *Communication in Statistics-Theory and Methods* 2004, **33**(8):1975-1991.

15. Manda SOM: **A nonparametric frailty model for clustered survival data**. *Communications in Statistics - Theory and Methods* 2011, **40**(5):863-875.

16. Hirsch K, Wienke A: **Software for semiparametric shared gamma and log-normal frailty models: An overview**. *Computer Methods And Programs In Biomedicine* 2012, **107**:582-597.

17. Duchateau L, Janssen P: **Penalized Partial Likelihood for Frailties and Smoothing Splines in Time to First Insemination Models for Dairy Cows**. *Biometrics* 2004, **60**:608-614.

18. Therneau TM, Grambsch PM, Pankratz VS: **Penalized survival models and frailty**. *12* 2003:156-175.

19. Ha ID, Noh M, Lee Y: **frailtyHL: A Package for Fitting Frailty Models with H-likelihood**. *The R Journal* 2012, **4**(2):28-37.

20. Ha ID, Lee Y, Song JK: **Hierarchical likelihood approach for frailty models**. *Biometrika* 2001, **88**:233-243.

21. Ha ID, Lee Y: **Comparison of hierarchical likelihood versus orthodox best linear unbiased predictor approaches for frailty models**. *Biometrika* 2005, **92**:717-723.

22. Martino S, Akerkar R, Rue H: **Approximate Bayesian Inference for Survival Models**. *Scandinavian Journal of Statistics* 2011, **38**:514-528.

23. Rue H, Martino S, Chopin N: **Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion)**. *J Roy Statist Soc Ser B* 2009, **71**:319-392.

24. Rondeau V, Commenges D, Joly P: **Maximum penalized likelihood estimation in a gamma-frailty model**. *Lifetime Data Analysis* 2003, **9**(2):139–153.

25. Joly P, Commenges D, Helmer C, Letenneur L: **A penalized likelihood approach for an illness-death model with interval-censored data: application to age-specific incidence of dementia**. *Biostatistics* 2002, **3**(3):433-443.

26. Monaco JV, Gorfine M, Hsu L: **General Semiparametric Shared Frailty Model: Estimation and Simulation with frailtySurv**. In.; 2017: 1-44.

Table 1: Mean-squared error of the parameter estimates

Parameter Estimates	INLA	frailtyHL	frailtyPack	coxph
Semiparametric Gamma Frailty				
Cluster size (n=50)				
Z1	0.0033	0.0036	0.0035	0.0035
Z2	0.0041	0.0046	0.0044	0.0044
Z3	0.0051	0.0054	0.0056	0.0058
Cluster size (n=100)				
Z1	0.0015	0.0017	0.0017	0.0017
Z2	0.0020	0.0022	0.0022	0.0022
Z3	0.0025	0.0026	0.0027	0.0027
Semiparametric Log-Normal Frailty				
Cluster size (n=50)				
Z1	0.0030	0.0033	0.0032	0.0032
Z2	0.0043	0.0047	0.0049	0.0046
Z3	0.0053	0.0058	0.0058	0.0058
Cluster size (n=100)				
Z1	0.0015	0.0016	0.0018	0.0016
Z2	0.0020	0.0022	0.0025	0.0022
Z3	0.0024	0.0026	0.0031	0.0026

NB: INLA- *Piecewise log-constant baseline hazard*; **frailtyHL**-*H-Likelihood*; **frailtyPack**-*Penalized full likelihood;* **CoxPH**-*Penalized partial likelihood*;

Table 2: Median and interquartile ranges of the frailty variances

Variable	INLA	FrailtyHL	FrailtyPack	CoxPh
Gamma frailty				
Gamma (cluster size n=50) frailty variance	1.1 (1.0-1.1)	1.9 (1.4-2.4)	1.8 (1.3-2.3)	1.8 (1.3-2.3)
Gamma (cluster size n=100) frailty variance	1.1 (1.0-1.1)	1.9 (1.4-2.5)	1.8 (1.3-2.3)	1.8 (1.3-2.3)
Log-Normal frailty				
Log-Normal (cluster size n=50) frailty variance	1.1 (1.1-1.2)	1.9 (1.3-2.5)	1.8 (1.2-2.5)	1.2 (0.9-1.6)
Log-Normal (cluster size n=100) frailty variance	1.1 (1.1-1.2)	1.8 (1.3-2.5)	1.7 (1.2-2.3)	1.2 (1.0-1.6)

NB: INLA- *Piecewise log-constant baseline hazard*; **frailtyHL**-*H-Likelihood*; **frailtyPack**-*Penalized full likelihood;* **CoxPH**-*Penalized partial likelihood*;

Figure

Click here to download Figure Figures.docx

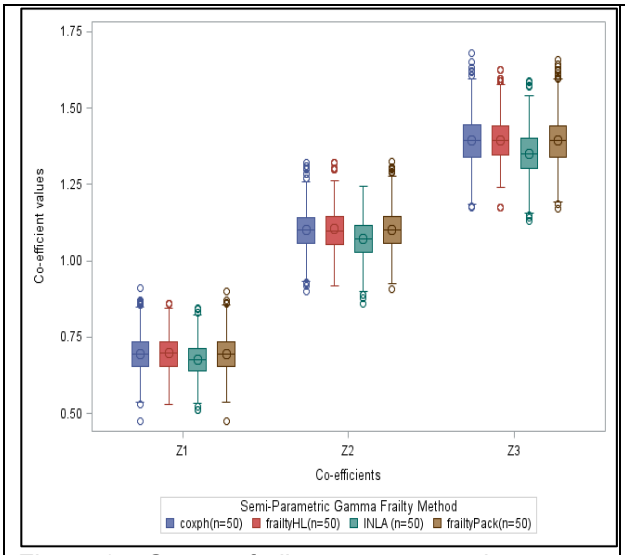


Figure 1a: Gamma frailty parameter estimates (Cluster size n=50)

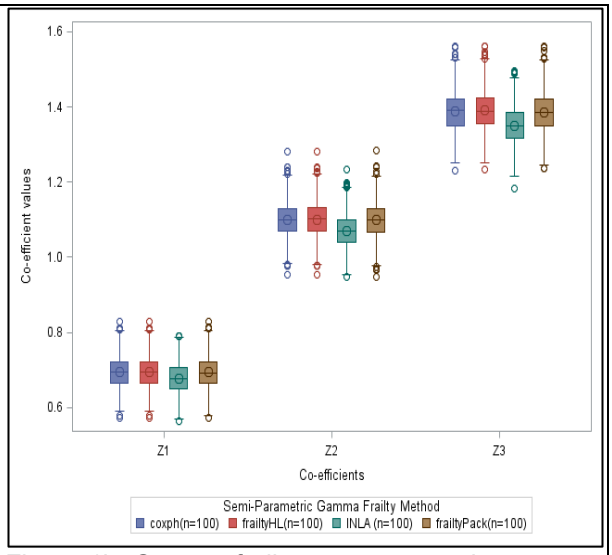


Figure 1b: Gamma frailty parameter estimates (Cluster size n=100)

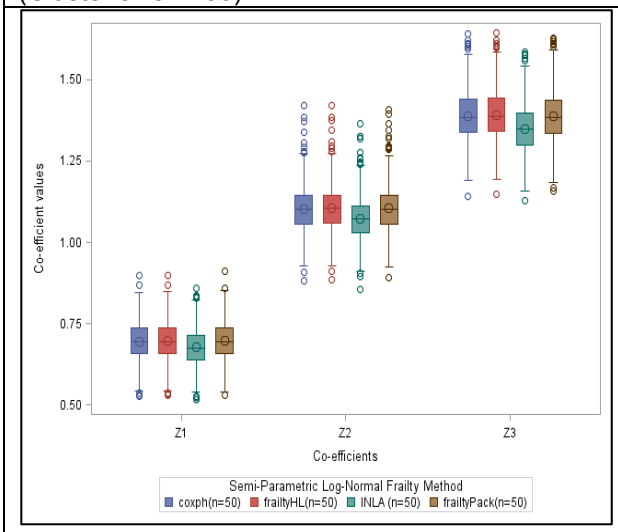


Figure 1c: Log-Normal frailty parameter estimates (Cluster size n=50)

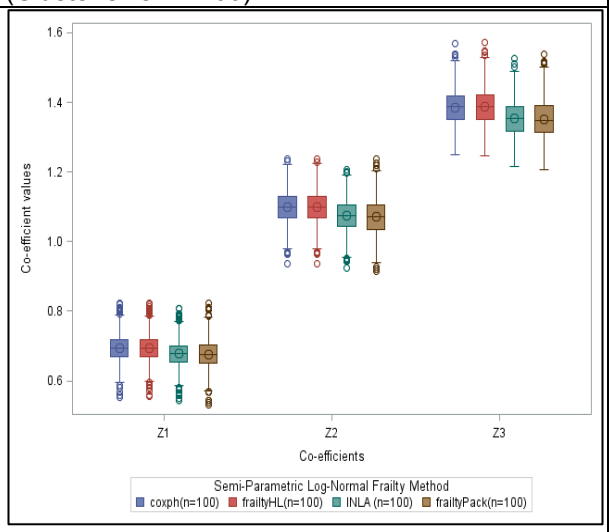


Figure 1d: Log-Normal frailty parameter estimates (Cluster size n=100)

Figure 1: Semiparametric gamma frailty parameter estimates from four methods

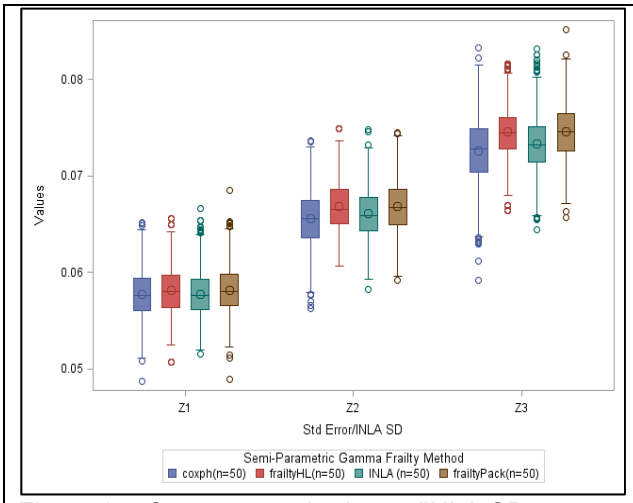


Figure 2a: Gamma standard error/INLA SD (cluster size n=50)

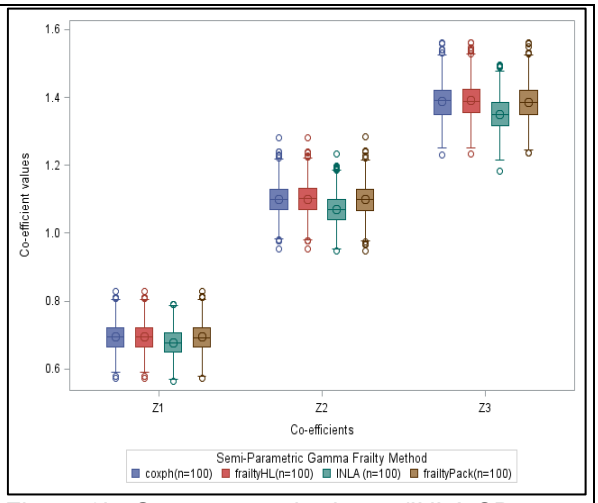


Figure 2b: Gamma standard error/INLA SD (Cluster size n=100)

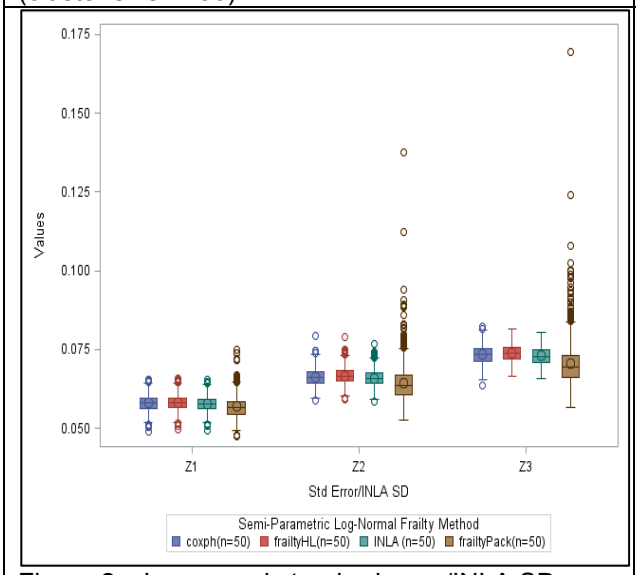


Figure 2c: Lognormal standard error/INLA SD (Cluster size n=50)

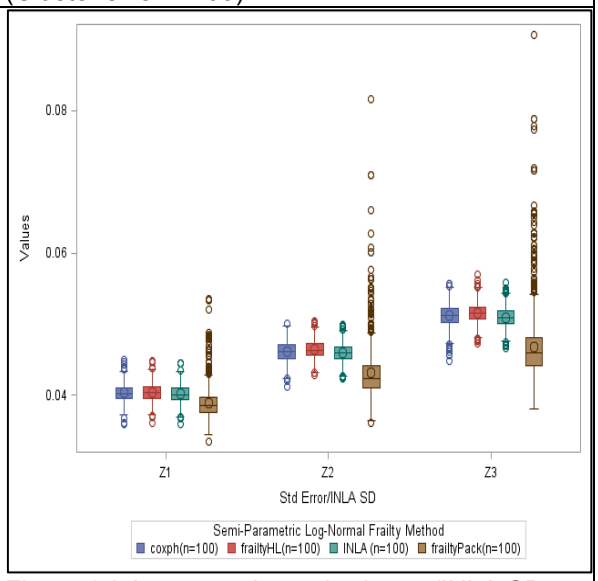


Figure 2d: Lognormal standard error/INLA SD (Cluster size n=100)

Figure 2: Semiparametric gamma frailty standard errors and INLA standard deviation from four methods

Global Health Action

ORIGINAL ARTICLE

Factors associated with mortality in HIV-infected people in rural and urban South Africa

Kennedy N. Otwombe[1,2]*, Max Petzold[3], Tebogo Modisenyane[1], Neil A. Martinson[1,4] and Tobias Chirwa[2]

[1]Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [2]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [3]Centre for Applied Biostatistics, Occupational and Environmental Medicine, University of Gothenburg, Gothenburg, Sweden; [4]Center for TB Research, Johns Hopkins University, Baltimore, MD, USA

Background: Factors associated with mortality in HIV-infected people in sub-Saharan Africa are widely reported. However rural–urban disparities and their association with all-cause mortality remain unclear. Furthermore, commonly used classical Cox regression ignores unmeasured variables and frailty.
Objective: To incorporate frailty in assessing factors associated with mortality in HIV-infected people in rural and urban South Africa.
Design: Using data from a prospective cohort following 6,690 HIV-infected participants from Soweto (urban) and Mpumalanga (rural) enrolled from 2003 to 2010; covariates of mortality were assessed by the integrated nested Laplace approximation method.
Results: We enrolled 2,221 (33%) rural and 4,469 (67%) urban participants of whom 1,555 (70%) and 3,480 (78%) were females respectively. Median age (IQR) was 36.4 (31.0–44.1) in rural and 32.7 (28.2–38.1) in the urban participants. The mortality rate per 100 person-years was 11 (9.7–12.5) and 4 (3.6–4.5) in the rural and urban participants, respectively. Compared to those not on HAART, rural participants had a reduced risk of mortality if on HAART for 6–12 (HR: 0.20, 95% CI: 0.10–0.39) and >12 months (HR: 0.10, 95% CI: 0.05–0.18). Relative to those not on HAART, urban participants had a lower risk if on HAART >12 months (HR: 0.35, 95% CI: 0.27–0.46).
The frailty variance was significant and >1 in rural participants indicating more heterogeneity. Similarly it was significant but <1 in the urban participants indicating less heterogeneity.
Conclusion: The frailty model findings suggest an elevated risk of mortality in rural participants relative to the urban participants potentially due to unmeasured variables that could be biological, socio–economic, or healthcare related. Use of robust methods that optimise data and account for unmeasured variables could be helpful in assessing the effect of unknown risk factors thus improving patient management and care in South Africa and elsewhere.

Keywords: *HIV*; *Mortality*; *Rural*; *Urban*; *Frailty*; *HAART*

Responsible Editor: Peter Byass, Umeå University, Sweden.

*Correspondence to: Kennedy N. Otwombe, Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, P.O Box 114, Diepkloof 1864, Johannesburg, South Africa, Email: otwombek@phru.co.za

Received: 18 July 2014; Revised: 28 August 2014; Accepted: 3 September 2014; Published: 29 September 2014

South Africa has the highest number of people living with HIV and the largest antiretroviral treatment (ART) programme globally (1). Driven by the need to achieve the millennium development goal of increasing the number of HIV-infected people accessing treatment (2), over 2 million people have been started on ART with another half million planned to be started each year. Mortality in HIV-infected individuals is high (3, 4), especially prior to ART treatment and while awaiting ART (5). This is associated with highly immunosuppressed patients with low CD4 counts and high viral loads. While survival on ART is markedly improved, mortality rates in the early stages of treatment remain relatively high (4, 6) signifying the importance of immunosuppression at the time of ART initiation. Knowledge of risk factors for mortality is important to improve ART programmes and to prolong life. Factors associated with mortality in HIV-infected people in sub-Saharan Africa are widely reported

Citation: Glob Health Action 2014, **7**: 25488 - http://dx.doi.org/10.3402/gha.v7.25488

with variations. They include CD4 and total lymphocyte counts, HIV RNA level, WHO and CDC staging, anaemia, malnutrition, and gender amongst others (6–11). However, prior studies do not specifically report and compare mortality in rural and urban settings. Furthermore time-to-event methods remain the most commonly used for the investigation of mortality risk factors particularly in HIV-infected cohorts (12). Cox proportional hazards regression analysis, the most popular approach assumes that the risk of mortality is homogeneous (13), an assumption that does not often hold (14). In epidemiological studies where some participants are related either genetically or are from the same family, share some unmeasured covariates, or have multiple recurrence events, their survival times are not independent due to within - subject dependency (15). Analysis of such data requires use of a method that accounts for dependence leading to heterogeneity such as frailty modelling (14–16), an extension of the Cox regression. Frailty modelling incorporates within-subject dependency by considering that the individual risk of events differs between subjects through accounting for unmeasured variables (15–17). Adjusting for within-subject correlation and accounting for unmeasured variables are some of the advantages of frailty modelling. Disadvantages include handling of competing risks and assuming the same risk within a cluster of unobserved variables.

Understanding the dynamics of mortality in HIV-infected people in sub-Saharan Africa is important as it influences patient management and care. As such, assessment of mortality surveillance patterns requires the use of robust statistical analysis methods that will optimise findings leading to appropriate policy decisions from evidence-based research. While current methods are well established, recent approaches, though more complex, provide further insight through their ability to model dependency. Detailed research from sub-Saharan Africa comparing factors associated with mortality in HIV-infected people between rural and urban areas using recently developed advanced methods are limited. Furthermore, those employing methods that account for unmeasured variables in this setting are not properly described. And, more importantly, we hypothesise that they inadequately capture uncertainty and are less precise. We therefore sought to assess and compare factors associated with mortality in HIV-infected adults in a large cohort of participants in a wellness programme in rural and urban South Africa while accounting for unmeasured variables in both sites.

Methods

Setting and study population

Data from two South African study sites were used: Perinatal HIV Research Unit (PHRU) in Soweto and Tintswalo hospital in Mpumalanga. In this study, participants enrolled in PHRU were classified as urban while those enrolled in Tintswalo were classified as rural. While both sites provided treatment using the same guideline from the department of health and similar resources such as nurses with similar qualifications, Tintswalo hospital is a primary health centre while the PHRU clinic is a research centre.

Soweto is an urban area in the city of Johannesburg in Gauteng, South Africa, with a population greater than a million people (18). It is in the southwestern part of Johannesburg where approximately 40% of the population resides. While many households are poor with high unemployment rates, there is a mix of poor and wealthy residents (18).

Tintswalo hospital is situated in the Bushbuckridge district of rural South Africa's northeast with a population of 541,247 people. Bushbuckridge is among the poorest areas of South Africa with up to 75% of the population living in poverty (19).

A nurse-based wellness programme was established in PHRU and Tinstwalo to provide pre- and post-treatment care for HIV-infected adults between the years 2003 and 2010. Participants in the study were referred from voluntary counselling and testing centres, surrounding clinics, hospital wards, or research programmes. Referrals for ART were made for those eligible, based on the South African ART treatment guidelines at the time (CD4 cell count <200 cells/mm^3) (20). Inclusion criteria were diagnosis of HIV, 18 years of age or older, receiving primary clinical care at PHRU or Tintswalo clinics and providing written informed consent. Further study procedures are presented elsewhere (21, 22).

Study questionnaires were administered at baseline and follow-up visits that occurred every 6 months. Losses to follow-up were minimised through actively following up participants by way of telephone calls, letters, and then home visits. Weight, CD4 counts, creatinine, and haemoglobin estimations were measured at scheduled six monthly visits while other investigations were done according to clinical presentation. To be considered positive for tuberculosis (TB), subjects had to meet at least one of the following criteria (21): initiation of multidrug TB therapy, presence of acid fast bacilli on microscopy or a biopsy suggestive of TB, mycobacterium culture positive for acid fast bacilli or *Mycobacterium tuberculosis*, hospitalisation due to TB or cause of death ascribed to TB.

Measures

For this study, the outcome measure was all-cause mortality, primarily through making contact with the next of kin. Socio-demographic measures included age, household income, number of people and rooms in a household, ever smoked and ever employed. Others were height and weight for calculating body mass index (BMI),

Citation: Glob Health Action 2014, **7**: 25488 - http://dx.doi.org/10.3402/gha.v7.25488

HAART status, ever had TB and CD4 counts. BMI was categorised into three groups using the CDC classification: underweight (<18.5), normal (18.5–24.9), or overweight/obese (≥ 25) while CD4 count was categorised into 0–200, 201–350, 351–500, and >500 cells/mm^3.

Ethics

Ethical approval was provided by the Human Research Ethics Committee of the University of the Witwatersrand.

Statistical analysis

Data were stratified by site. Continuous measures at enrolment such as age, BMI, and CD4 count were assessed descriptively by median and interquartile ranges and compared between the two groups using the Kruskal–Wallis non-parametric test. Frequencies and associated proportions were determined for categorical measures and compared using the Fishers Exact and Chi–square analysis as appropriate.

Mortality rates (95% confidence interval) per 100 person-years were determined using Poisson regression modelling by site for different lengths of time pre- and post-HAART and presented graphically. The Kaplan–Meier test was used to determine differences in surveillance patterns by site.

Factors associated with mortality by site were assessed using the integrated nested Laplace approximation (INLA) (23, 24), where univariate and multivariate shared gamma frailty models were fitted. INLA is a Bayesian inference technique, computationally efficient due to its speed relative to existing methods and more accurate parameter estimates. For the purpose of interpretation, posterior Bayesian estimates generated from INLA were exponentiated to determine their approximate hazard ratios and 95% confidence intervals (25).

INLA accounted for unmeasured variables and allowed for the comparison of frailty variances between sites. A significant frailty variance with a value >1 suggests a higher rate for the event (or shorter survival times) than would be predicted under the basic Cox model while <1 suggests a lower rate (longer survival times). Hence, failure to account for frailty may overestimate or underestimate the hazard rate.

Backward selection together with inclusion of plausible socio-demographic and clinical factors was used to determine the final multivariate model. Socio-demographic variables included gender, ever employed, and ever smoked. Clinical variables included HAART status, BMI, CD4 cell count, and ever had TB. The most recent information was used for HAART status while BMI and CD4 cell count values were considered over time.

A 95% confidence interval excluding one was determined as significant. Model fit was assessed using the deviance information criterion.

Multiple imputation (MI) under the missing-at-random mechanism was used to impute 18% (n = 1,195) and 11% (n = 704) of missing CD4 count and BMI data, respectively (26–28). Five imputed datasets reflecting different scenarios were generated and analysed following the MI procedure.

Several sensitivity analyses were performed using INLA. This included assessing factors associated with mortality overall, pre- and post-HAART. Kaplan–Meier plots of time to death by gender, HAART use, CD4 count, WHO staging, and BMI were also fitted. For participants that were lost to follow-up, their time in the study up to their last visit was included in the analysis.

All statistical analyses were performed using R and SAS Enterprise Guide 5.1 under the assumption of a two-sided test at 5% significance level.

Results

Overall, 6,690 participants were enrolled and followed up for a cumulative total of 2225.8 and 8185.9 person-years in the rural (33%) and urban (67%) sites, respectively. The median (IQR) pre- and post-HAART time in years was 0.6 (0–1.9) and 1.1 (0.5–2.6), respectively. Overall, 75% were women: rural (70%) and urban (78%). The median age at enrolment was 36.4 (IQR: 31.0–44.1) and 32.7 (IQR: 28.2–38.1) years for rural and urban sites, respectively.

Rural participants were more likely to report a household income of $\leq$R1,000 per month compared to those in the urban areas (72.8% vs. 48.5%; $p<0.0001$). The proportion not on HAART at the end of the study was significantly higher in the urban area compared to the rural area (63.2% vs. 30.0%; $p<0.0001$) while unemployment was significantly higher in the rural area (86.7% vs. 70.7%; $p<0.0001$) relative to the urban areas.

At enrolment, 54% of the rural and 27% of the urban participants had either started HAART or were already on HAART. On follow-up, 17 and 10% of rural and urban participants, respectively, initiated HAART.

The median BMI of urban participants was higher than that of the rural participants (24.9 vs. 21.0; $p<0.0001$). Similarly, the median CD4 count at entry of rural participants was lower than that in the urban participants (178 cells/mm^3 vs. 303 cells/mm^3; $p<0.0001$, Table 1). The proportion on HAART for <6 months and with a most recent CD4 count <200 cells/mm^3 was significantly higher in the urban compared to the rural participants (76% vs. 59%; $p<0.0001$). Similarly, urban participants on HAART for 6–12 months and with a most recent CD4 count <200 cells/mm^3 were significantly more than in the rural participants (69% vs. 53%; $p=0.0004$). The proportion unemployed and with a CD4 count <200 cells/mm^3 was significantly higher in the rural participants compared to the urban participants (40% vs. 24%; $p<0.0001$). The proportion ever with TB in the rural participants was significantly higher than in the urban participants (44.8% vs. 25.4%; $p<0.0001$, Table 2).

Table 1. Participant characteristics at enrolment by site

	Rural			Urban		
Variable	Overall (n = 2,221)	Male (n = 666, 30%)	Female (n = 1,555, 70%)	Overall (n = 4,469)	Male (n = 989, 22%)	Female (n = 3,480, 78%)
Median age (IQR) in years	36.4 (31.0–44.1)	39.2 (33.3–48.1)	35.2 (30.0–42.5)	32.7 (28.2–38.1)	35.7 (31.5–41.5)	31.7 (27.4–37.0)
Age group in years						
18–25 (%)	216 (9.7)	30 (4.5)	186 (12.0)	672 (15.0)	65 (6.6)	607 (17.4)
26–30 (%)	342 (15.4)	72 (10.8)	270 (17.4)	1,117 (25.0)	158 (16.0)	959 (27.6)
31–34 (%)	420 (18.9)	111 (16.7)	309 (19.9)	984 (22.0)	222 (22.4)	762 (21.9)
35–39 (%)	424 (19.1)	136 (20.4)	288 (18.5)	864 (19.3)	250 (25.3)	614 (17.6)
>40 (%)	819 (36.9)	317 (47.6)	502 (32.3)	832 (18.6)	294 (29.7)	538 (15.5)
Median CD4 count (IQR)	178.0 (81.0–333.6)	162.0 (69.0–289.9)	184.0 (86.0–350.6)	303.0 (156.0–470.7)	281.0 (142.0–443.7)	309.0 (163.0–478.2)
CD4 categories						
0–200 (%)	1,235 (55.6)	398 (59.8)	837 (53.8)	1,474 (33.0)	367 (37.1)	1,107 (31.8)
201–350 (%)	472 (21.3)	144 (21.6)	328 (21.1)	1,102 (24.7)	241 (24.4)	861 (24.7)
351–500 (%)	284 (12.8)	75 (11.3)	209 (13.4)	915 (20.5)	190 (19.2)	725 (20.8)
>500 (%)	230 (10.4)	49 (7.4)	181 (11.6)	978 (21.9)	191 (19.3)	787 (22.6)
Median BMI (IQR)	21.0 (17.9–25.0)	19.9 (17.2–23.4)	21.4 (18.4–25.8)	24.9 (20.7–29.2)	22.5 (18.2–26.4)	25.6 (21.4–30.0)
BMI						
Underweight (%)	638 (29.0)	241 (36.9)	397 (25.7)	594 (13.4)	240 (24.8)	354 (10.2)
Normal (%)	987 (44.9)	299 (45.7)	688 (44.6)	1,612 (36.4)	393 (40.6)	1,219 (35.2)
Obese/overweight (%)	572 (26.0)	114 (17.4)	458 (29.7)	2,226 (50.2)	336 (34.7)	1,890 (54.6)
Median crowding index (IQR)	1.5 (1.0–2.3)	1.3 (0.8–2.0)	1.6 (1.0–2.5)	1.5 (1.0–2.3)	1.0 (1.0–2.0)	1.5 (1.0–2.5)
Median household income (IQR) in Rands	600.0 (189.8–1010.0)	780.0 (50.0–1500.0)	500.0 (190.0–974.7)	1080.0 (500.0–2130.0)	1200.0 (600.0–2400.0)	1020.0 (500.0–2020.0)
Household income						
≤R1,000 (%)	1,516 (72.8)	423 (67.4)	1,093 (75.2)	2,161 (48.5)	453 (45.9)	1,708 (49.2)
R1,001–R5,000 (%)	528 (25.4)	193 (30.7)	335 (23.1)	2,092 (46.9)	487 (49.3)	1,605 (46.3)
>R5,000 (%)	37 (1.8)	12 (1.9)	25 (1.7)	203 (4.6)	47 (4.8)	156 (4.5)

BMI, body mass index; HAART, highly active antiretroviral therapy; TB, tuberculosis.

Crowding index is the number of people divided by the number of rooms in a household; household income refers to household monthly income; for household income, 1USD was equivalent to R7 at the time of the study.

Citation: Glob Health Action 2014, **7**: 25488 - http://dx.doi.org/10.3402/gha.v7.25488

Table 2. Distribution of HAART status, employment, smoking, and TB by site

Variable	Rural			Urban		
	Overall (n = 2,221)	Male (n = 666, 30%)	Female (n = 1,555, 70%)	Overall (n = 4,469)	Male (n = 989, 22%)	Female (n = 3,480, 78%)
Time on HAART						
No HAART (%)	666 (30.0)	174 (26.1)	492 (31.6)	2,826 (63.2)	555 (56.1)	2,271 (65.3)
<6 months (%)	468 (21.0)	153 (23.0)	315 (20.3)	356 (8.0)	108 (10.9)	248 (7.1)
6–12 months (%)	434 (20.0)	133 (20.0)	301 (19.4)	178 (4.0)	35 (3.5)	143 (4.1)
>12 months (%)	653 (29.0)	206 (30.9)	447 (28.7)	1,109 (24.8)	291 (29.5)	818 (23.5)
Ever employed						
Yes (%)	296 (13.3)	143 (21.5)	153 (9.8)	1,309 (29.3)	400 (40.4)	909 (26.1)
No (%)	1,925 (86.7)	523 (78.5)	1,402 (90.2)	3,160 (70.7)	589 (59.6)	2,571 (73.9)
Ever smoked						
Yes (%)	358 (16.1)	295 (44.3)	63 (4.1)	1,284 (28.7)	706 (71.4)	578 (16.6)
No (%)	1,863 (83.9)	371 (55.7)	1,492 (95.9)	3,185 (71.3)	283 (28.6)	2,902 (83.4)
Ever had TB						
Yes (%)	995 (44.8)	350 (52.6)	645 (41.5)	1,134 (25.4)	313 (31.6)	821 (23.6)
No (%)	1,226 (55.2)	316 (47.4)	910 (58.5)	3,335 (74.6)	676 (68.4)	2,659 (76.4)

The most recent value is provided for time on HAART.

Overall, 566 (8%) of the participants died during follow-up: 11 and 7% in rural and urban sites, respectively ($p < 0.0001$). Of these, majority were females (63% rural vs. 69% urban; $p = 0.1952$). Overall, the mortality rate per 100 person-years in rural sites was higher at 11 (95% CI: 9.7–12.5) compared to urban sites at 4 (95% CI: 3.6–4.5). In the first 6 months, 48% of rural participants had died while 51% of the deaths in the urban site occurred after 12 months. Mortality rates per 100 person-years while on HAART were 11 (95% CI: 8.7–14.0) and 7 (95% CI: 5.9–8.3) for rural and urban participants, respectively. When taking HAART, majority of the deaths in the rural site occurred within the first 3 months of treatment initiation (49%) whereas they occurred after 12 months in the urban site (44%). Further description of those who died pre- and post-treatment in the two sites by CD4 category and BMI is presented in Table 3. Relative to the urban participants, mortality rates in the rural participants was higher (Fig. 1).

In the rural participants, being on HAART for 6–12 months (HR: 0.20, 95% CI: 0.10–0.39) and >12 months (HR: 0.10, 95% CI: 0.05–0.18) relative to no HAART was protective after adjusting for CD4 count (Table 4). However, underweight BMI compared to normal BMI was associated with a higher risk of mortality (HR: 1.74, 95% CI: 1.06–2.85).

Being on HAART for <6 months (HR: 4.48, 95% C: 3.0–6.56) and 6–12 months (HR: 2.39, 95% CI: 1.60–3.47) compared to no HAART was associated with an elevated risk of mortality in the urban participants after adjusting for CD4 count, employment, smoking, and TB history. However, those with overweight/obese BMI (HR: 0.73, 95% CI: 0.56–0.90) had a significantly lower risk of mortality compared to those with normal BMI.

The estimated frailty variances were statistically significant for rural and urban sites; 5.65 (95% CI: 4.50–7.19) and 5.4×10^{-5} (95% CI: $1.5 \times 10^{-5} - 8.0 \times 10^{-4}$) respectively.

Discussion

In sub-Saharan Africa where the burden of HIV is the greatest, research on factors associated with all-cause mortality in HIV-infected cohorts has widely been done. However, research on risk factors using advanced Bayesian methods such as INLA that incorporate unmeasured variables is limited (23, 24). Common protective factors against mortality included HAART use for >12 months and CD4 count above 200 cells/mm^3 while underweight BMI was positively associated with mortality. Even with differences between the rural and urban participants, the hazard rates were often in the same direction.

While both frailty variances were significant, our findings show that the frailty variance for participants in the rural site was higher than that of the urban. This suggests important unmeasured variables relevant to participants in the rural site may not have been considered (as the variance was >1) signifying more heterogeneity (variability). In contrast, the variance for participants in the urban site was <1 signifying less heterogeneity and probably that the model accounted for most of the important variables. Available literature on covariates of mortality relies on classical statistical methods that do

Table 3. Characteristics of those who died

	Rural			Urban		
Variable	Overall	Male	Female	Overall	Male	Female
Total deaths	242	88	154	324	101	223
Mortality rate per 100 person-years (95% CI)	11 (9.7–12.5)	14 (11.4–17.3)	10 (8.5–11.7)	4 (3.6–4.5)	6 (4.9–7.3)	3 (2.6–3.4)
Time to death in months						
0–3 (%)	76 (31.5)	30 (34.1)	46 (30.1)	38 (12.1)	17 (17.7)	21 (9.6)
3–6 (%)	40 (16.6)	16 (18.2)	24 (15.7)	47 (15.0)	16 (16.7)	31 (14.2)
6–9 (%)	32 (13.3)	10 (11.4)	22 (14.4)	36 (11.5)	9 (9.4)	27 (12.4)
9–12 (%)	26 (10.8)	10 (11.4)	16 (10.5)	27 (8.6)	5 (5.2)	22 (10.1)
>12 (%)	67 (27.8)	22 (25)	45 (29.4)	166 (52.9)	49 (51)	117 (53.7)
CD4 categories						
0–200 (%)	167 (69.0)	59 (67.0)	108 (70.1)	195 (60.2)	66 (65.3)	129 (57.8)
201–350 (%)	35 (14.5)	11 (12.5)	24 (15.6)	69 (21.3)	18 (17.8)	51 (22.9)
351–500 (%)	24 (9.9)	11 (12.5)	13 (8.4)	36 (11.1)	11 (10.9)	25 (11.2)
>500 (%)	16 (6.6)	7 (8.0)	9 (5.8)	24 (7.4)	6 (5.9)	18 (8.1)
BMI						
Underweight (%)	109 (45.2)	41 (46.6)	68 (44.7)	75 (23.6)	34 (35.1)	41 (18.6)
Normal (%)	95 (39.6)	37 (42.0)	58 (38.2)	126 (39.6)	40 (41.2)	86 (38.9)
Overweight/obese (%)	36 (15.0)	10 (11.4)	26 (17.1)	117 (36.8)	23 (23.7)	94 (42.5)
Deaths on HAART	68	31	37	133	52	81
Mortality rate per 100 person-years (95% CI)	11 (8.7–14.0)	15 (10.5–21.3)	9 (6.5–12.4)	7 (5.9–8.3)	9 (6.9–11.8)	6 (4.8–7.5)
Time to death in months						
0–3 (%)	33 (48.5)	15 (48.4)	18 (48.6)	21 (16.4)	10 (20.4)	11 (13.9)
3–6 (%)	12 (7.6)	5 (16.1)	7 (18.9)	17 (13.3)	8 (16.3)	9 (11.4)
6–9 (%)	9 (13.2)	3 (9.7)	6 (16.2)	18 (14.1)	6 (12.2)	12 (15.2)
9–12 (%)	4 (5.9)	2 (6.5)	2 (5.4)	13 (10.2)	2 (4.1)	11 (13.9)
>12 (%)	10 (14.7)	6 (19.4)	4 (10.8)	59 (46.1)	23 (46.9)	36 (45.6)
CD4 categories						
0–200 (%)	43 (63.2)	19 (16.3)	24 (64.9)	86 (64.7)	39 (75.0)	47 (58.0)
201–350 (%)	11 (6.2)	3 (9.7)	8 (21.6)	23 (17.3)	6 (11.5)	17 (21.0)
351–500 (%)	9 (13.2)	6 (19.4)	3 (8.1)	13 (9.8)	3 (5.8)	10 (12.3)
>500 (%)	5 (7.4)	3 (9.7)	2 (5.4)	11 (8.3)	4 (7.7)	7 (8.6)
BMI						
Underweight (%)	25 (37.3)	13 (41.9)	12 (33.3)	30 (22.9)	19 (38.0)	11 (13.6)
Normal (%)	24 (35.8)	9 (29.0)	15 (41.7)	48 (36.6)	21 (42.0)	27 (33.3)
Overweight/Obese (%)	18 (26.9)	9 (29.0)	9 (25.0)	53 (40.5)	10 (20.0)	43 (53.1)

not account for unmeasured variables (12). Potential unmeasured variables include 1) biological factors such as undiagnosed TB and/or other opportunistic infections such as cryptococcosis, 2) socio-economic factors that were not included in the data collection process, 3) healthcare-related factors, 4) adherence to HAART medication, and 5) cultural and religious beliefs. A study from Portugal assessed predictors of mortality in HIV-associated hospitalisations, and through frailty modelling, it was shown how unmeasured variables in the form of quality in healthcare in different hospitals affected mortality (17). Hence, an assessment that allows for the use of a mixed methods approach may provide an opportunity for further insight of the unmeasured variables that may qualify as risk factors.

Underweight BMI was associated with mortality in both rural and urban areas with a slightly higher hazard in the urban group. Our findings concur with others from previous studies (5, 21, 29). While overweight or obese BMI was not associated with mortality in rural participants, it was associated with decreased mortality in the urban participants. Previous studies have shown that overweight or obese BMI reduces the risk of mortality (21). The proportion of participants with underweight BMI was higher in the rural group while that of overweight or obese participants was higher in the urban group

Citation: Glob Health Action 2014, **7**: 25488 - http://dx.doi.org/10.3402/gha.v7.25488

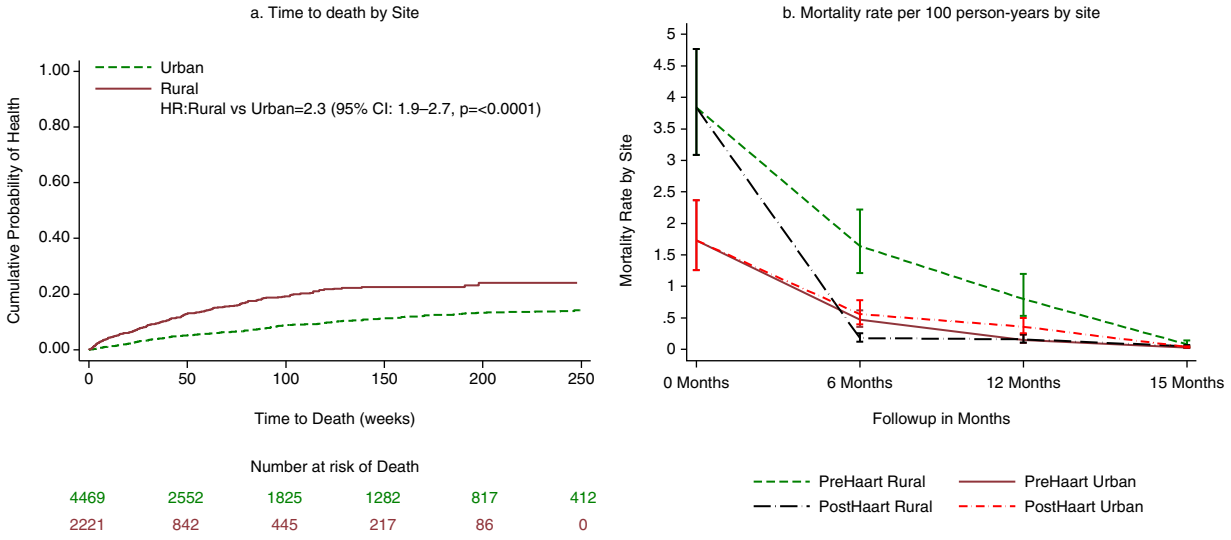


Fig. 1. Survival and mortality rate plots by site. a) displays the cumulative probability of death by site in the cohort of HIV-infected people in South Africa. b) is a plot of mortality rates during follow-up by site showing the period prior to initiating HAART to post-HAART.

suggesting differences in nutrition between the two groups. It may also be that the use of HAART mediated in the lack of association between mortality and overweight or obese BMI in the rural participants.

Higher, most recent CD4 count was protective in both groups as similarly reported in previous studies (30, 31). However at enrolment, rural participants had a lower median CD4 count that qualified them to initiate therapy

Table 4. Factors associated with mortality in rural and urban South Africa

	Rural		Urban	
	Univariate	Multivariate	Univariate	Multivariate
Variable	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)
Gender				
Female vs. male	0.69 (0.45–1.1)	0.78 (0.48–1.27)	0.53 (0.42–0.67)	1.09 (0.80–1.44)
Time on HAART				
<6 months vs. No HAART	1.12 (0.56–2.22)	1.16 (0.62–2.16)	17.25 (11.74–24.78)	**4.48 (3.0–6.56)**
6–12 months vs. No HAART	0.17 (0.08–0.37)	**0.20 (0.10–0.39)**	7.20 (4.89–10.35)	**2.39 (1.60–3.47)**
>12 months vs. No HAART	0.10 (0.05–0.20)	**0.10 (0.05–0.18)**	1.03 (0.80–1.32)	**0.35 (0.27–0.46)**
BMI				
Underweight vs. normal	2.20 (1.40–3.24)	**1.74 (1.06–2.85)**	4.29 (3.26–5.62)	**2.49 (1.87–3.11)**
Overweight/obese vs. normal	0.55 (0.30–0.98)	0.61 (0.31–1.16)	0.55 (0.42–0.71)	**0.73 (0.56–0.90)**
CD4 Count (cells/mm^3)				
201–350 vs. 0–200	0.26 (0.11–0.56)	**0.25 (0.13–0.47)**	0.19 (0.14–0.24)	**0.22 (0.16–0.29)**
351–500 vs. 0–200	0.49 (0.20–1.08)	**0.44 (0.21–0.89)**	0.09 (0.06–0.13)	**0.11 (0.07–0.17)**
>500 vs. 0–200	0.25 (0.08–0.66)	**0.22 (0.09–0.53)**	0.04 (0.02–0.08)	**0.05 (0.03–0.09)**
Ever employed				
No vs. Yes	0.95 (0.53–1.83)	1.13 (0.58–2.24)	2.03 (1.55–2.70)	**1.87 (1.42–2.48)**
Ever smoked				
Yes vs. No	0.99 (0.56–1.56)		2.07 (1.66–2.59)	**1.37 (1.04–1.79)**
Ever had TB				
Yes vs. No	1.66 (1.10–2.47)	1.55 (0.98–2.46)	2.84 (2.29–3.54)	**2.76 (2.19–3.48)**

Bold entries represent significant values.

as per the treatment guidelines at the time. As in many treatment programmes across sub-Saharan Africa, first time testers often present for care with low CD4 count. Previous studies have underscored the relationship between treatment initiation, low CD4 count, and early mortality (32–35).

Those on treatment for >6 months had a lower risk of mortality in the rural participants compared to >12 months in the urban participants. Among those on treatment for 6–12 months in the urban participants, a greater proportion had a most recent CD4 count <200 cells/mm^3. Findings from prior research show an elevated risk of early mortality especially in immunocompromised patients in the early period of treatment (4, 9, 34, 36, 37). The high mortality rate in urban participants may be due to late initiation of treatment leading to advanced HIV disease. Furthermore, the association between late initiation of treatment and low CD4 count is well established. It may also be that mortality was accelerated in this group by opportunistic infections that were not identified in time including undiagnosed and untreated TB. TB remains the most common opportunistic infection diagnosed in HIV-infected people.

Unemployed urban participants had a higher risk of mortality, a finding that concurs with previous research (38–41). HIV care was provided at no cost in both sites irrespective of their employment status. It maybe that mortality is affected by differential access to care in urban participants. Other studies have shown that mortality could be affected by poor social situations such as unemployment, poor housing, and social isolation. Furthermore, these factors may influence one's health negatively (38, 39).

There was an association between ever suffering from TB and mortality in participants from the urban site. Contrary to our findings, a previous study from Durban in South Africa found no association (36). TB is a common opportunistic infection in HIV-infected individuals and its association with mortality is indisputable (21, 37, 42). Since urban participants were significantly more likely to have CD4 counts <200 cells/mm^3 up to 12 months on HAART, it may be that they are immune-compromised and more susceptible to opportunistic infections, such as TB, and early mortality.

Some variables such as employment were self-reported and may have been influenced by social desirability bias. Using cumulative variables such as 'ever smoking' or 'ever employed' may lead to misclassification given the fact that it is the cumulative, rather than the instantaneous. While all attempts were made to collect and report all measures, there were some missing values that were estimated by MIs which may have introduced some bias. Some deaths may have been misclassified since some participants were lost to follow-up and reasons for loss to follow-up were not systematically recorded. However, an advantage of our study was the large sample size and the use of a Bayesian frailty survival model that accounts for unmeasured factors.

Conclusion

Our findings suggest that using robust advanced statistical analysis methods may provide further insight to risk factors for mortality. Furthermore, and where appropriate, a mixture of methods involving both quantitative and qualitative approaches could be employed allowing for further understanding of the unmeasured variables. Such an approach is likely to inform HIV care managers and policy makers of important factors for which interventions could be developed or existing ones enhanced to improve patient management and care. Any such interventions should alleviate rural–urban healthcare disparities not only in South Africa but also in other countries with similar experiences. Interventions should specifically help in alleviating the challenges faced by rural communities in accessing care.

Acknowledgements

We are grateful to all the participants who enrolled in this study and allowed their data to be used for research purpose, the clinical staff for providing care and treatment, and the Statistical and Data Management Center at the Perinatal HIV Research Unit.

Financial support

Patient care was funded by a US PEPFAR grant through USAID South Africa (674-A-00-08-00009-00).

Analysis was supported by the Consortium for Advanced Research Training in Africa (CARTA). CARTA is jointly led by the African Population and Health Research Center and the University of the Witwatersrand and funded by the Wellcome Trust (UK) (Grant No: 087547/Z/08/Z), the Department for International Development (DfID) under the Development Partnerships in Higher Education (DelPHE), the Carnegie Corporation of New York (Grant No: B 8606), the Ford Foundation (Grant No: 1100-0399), Google.Org (Grant No: 191994), Sida (Grant No: 54100029), and MacArthur Foundation Grant No: 10-95915-000-INP.

KO received an academic scholarship from Canadian–Africa Prevention Trials Network and the Forgaty International Center (grant no. TW007370/3).

Conflict of interest and funding

The authors have not received any funding or benefits from industry or elsewhere to conduct this study.

References

1. UNAIDS (2010). Outlook. Available from: http://www.unaids.org/en/resources/presscentre/featurestories/2010/july/20100713outloklaunch/ [cited 27 February 2014].

2. UN (2013). The millenium development goals report. Available from: http://www.un.org/millenniumgoals/pdf/report-2013/mdg-report-2013-english.pdf [cited 27 February 2014].
3. Oliveira I, Andersen A, Furtado A, Medina C, da Silva D, da Silva ZJ, et al. Assessment of simple risk markers for early mortality among HIV-infected patients in Guinea-Bissau: a cohort study. BMJ Open 2012; 2. 1–10.
4. Otwombe KN, Laher F, Tutu-Gxashe T, Gray G, Mohapi L. The effect of a maturing antiretroviral program on early mortality for patients with advanced immune-suppression in Soweto, South Africa. PLoS One 2013; 8: e81538.
5. Masiira B, Baisley K, Mayanja BN, Kazooba P, Maher D, Kaleebu P. Mortality and its predictors among antiretroviral therapy naive HIV-infected individuals with CD4 cell count (350 cells/mm3 compared to the general population: data from a population-based prospective HIV cohort in Uganda. Glob Health Action 2014; 7: 21843.
6. Fetzer BC, Hosseinipour MC, Kamthuzi P, Hyde L, Bramson B, Jobarteh K, et al. Predictors for mortality and loss to follow-up among children receiving anti-retroviral therapy in Lilongwe, Malawi. Trop Med Int Health 2009; 14: 862–9.
7. Badri M, Lawn SD, Wood R. Short-term risk of AIDS or death in people infected with HIV-1 before antiretroviral therapy in South Africa: a longitudinal study. Lancet 2006; 368: 1254–9.
8. Erikstrup C, Kallestrup P, Zinyama R, Gomo E, Mudenge B, Gerstoft J, et al. Predictors of mortality in a cohort of HIV-1-infected adults in rural Africa. J Acquir Immune Defic Syndr 2007; 44: 478–83.
9. Jerene D, Endale A, Hailu Y, Lindtjorn B. Predictors of early death in a cohort of Ethiopian patients treated with HAART. BMC Infect Dis 2006; 6: 136.
10. Johannessen A, Naman E, Ngowi BJ, Sandvik L, Matee MI, Aglen HE, et al. Predictors of mortality in HIV-infected patients starting antiretroviral therapy in a rural hospital in Tanzania. BMC Infect Dis 2008; 8: 52.
11. Kouanda S, Meda IB, Nikiema L, Tiendrebeogo S, Doulougou B, Kabore I, et al. Determinants and causes of mortality in HIV-infected patients receiving antiretroviral therapy in Burkina Faso: a five-year retrospective cohort study. AIDS Care 2012; 24: 478–90.
12. Otwombe KN, Petzold M, Martinson N, Chirwa T. A review of the study designs and statistical methods used in the determination of predictors of all-cause mortality in HIV-infected cohorts: 2002–2011. PLoS One 2014; 9: e87356.
13. Cox D. Regression models and life-tables. J Roy Stat Soc B 1972; 43: 187–220.
14. Chuang SK, Cai T, Douglass CW, Wei LJ, Dodson TB. Frailty approach for the analysis of clustered failure time observations in dental research. J Dent Res 2005; 84: 54–8.
15. Wienke A. Frailty models in survival analysis. Florida, USA: Chapman and Hall/CRC; 2011.
16. Vaupel JW, Manton KG, Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 1979; 16: 439–54.
17. Dias SS, Andreozzi V, Martins MO, Torgal J. Predictors of mortality in HIV-associated hospitalizations in Portugal: a hierarchical survival model. BMC Health Serv Res 2009; 9: 125.
18. City of Johannesburg (2012). 2012/16 Integrated Development Plan. Available from: http://www.joburg-archive.co.za/2012/pdfs/idp/idp201216.pdf [cited 6 August 2014].
19. Bushbuckridge Municipality (2010). Bushbuckridge local municipality local development strategy. Available from: http://bushbuckridge.gov.za/wp-content/uploads/2014/03/Bushbuckridge-LED-Strategy-2010–20142.pdf [cited 11 June 2014].
20. Southern African HIV Clinicians Society: antiretroviral therapy in adults. South Afr J HIV Med 2008; 9: 18–31.
21. Hanrahan CF, Golub JE, Mohapi L, Tshabangu N, Modisenyane T, Chaisson RE, et al. Body mass index and risk of tuberculosis and death. AIDS 2010; 24: 1501–8.
22. Venkatesh KK, de Bruyn G, Lurie MN, Mohapi L, Pronyk P, Moshabela M, et al. Decreased sexual risk behavior in the era of HAART among HIV-infected urban and rural South Africans attending primary care clinics. AIDS 2010; 24: 2687–96.
23. Martino S, Akerkar R, Rue H. Approximate Bayesian inference for survival models. Scand J Stat 2010; 38: 514–28.
24. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J Roy Stat Soc B 2009; 71: 319–92.
25. Fairley L, Forman D, West R, Manda S. Spatial variation in prostate cancer survival in the Northern and Yorkshire region of England using Bayesian relative survival smoothing. Br J Canc 2008; 99: 1786–93.
26. Little RJA. Regression with missing X's: a review. J Am Stat Assoc 1992; 87: 1227–37.
27. Rubin DB. Multiple imputation after 18+ years. J Am Stat Assoc 1996; 91: 473–89.
28. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med 1991; 10: 585–98.
29. Koethe JR, Blevins M, Nyirenda CK, Kabagambe EK, Chiasera JM, Shepherd BE, et al. Serum phosphate predicts early mortality among underweight adults starting ART in Zambia: a novel context for refeeding syndrome? J Nutr Metabol 2013; 2013: 545439.
30. Poka-Mayap V, Pefura-Yone EW, Kengne AP, Kuaban C. Mortality and its determinants among patients infected with HIV-1 on antiretroviral therapy in a referral centre in Yaounde, Cameroon: a retrospective cohort study. BMJ Open 2013; 3. 1–7.
31. Tadesse K, Haile F, Hiruy N. Predictors of mortality among patients enrolled on antiretroviral therapy in Aksum Hospital, Northern Ethiopia: a retrospective cohort study. PLoS One 2014; 9: e87392.
32. Barth RE, Tempelman HA, Moraba R, Hoepelman AIM. Long-term outcome of an HIV-treatment programme in rural Africa: viral suppression despite early mortality. AIDS Res Treat 2011; 2011: 434375.
33. Mageda K, Leyna GH, Mmbaga EJ. High initial HIV/AIDS-related mortality and -its predictors among patients on anti-retroviral therapy in the Kagera Region of Tanzania: a five-year retrospective cohort study. AIDS Res Treat 2012; 2012: 843598.
34. Schmaltz CA, Santoro-Lopes G, Lourenco MC, Morgado MG, Velasque Lde S, Rolla VC. Factors impacting early mortality in tuberculosis/HIV patients: differences between subjects naive to and previously started on HAART. PLoS One 2012; 7: e45704.
35. Wubshet M, Berhane Y, Worku A, Kebede Y, Diro E. High loss to followup and early mortality create substantial reduction in patient retention at antiretroviral treatment program in north-west Ethiopia. ISRN AIDS 2012; 2012: 721720.
36. Ojikutu BO, Zheng H, Walensky RP, Lu Z, Losina E, Giddy J, et al. Predictors of mortality in patients initiating antiretroviral therapy in Durban, South Africa. S Afr Med J 2008; 98: 204–8.
37. Russell EC, Charalambous S, Pemba L, Churchyard GJ, Grant AD, Fielding KL. Low haemoglobin predicts early mortality among adults starting antiretroviral therapy in an HIV care programme in South Africa: a cohort study. BMC Public Health 2010; 10: 433.
38. Burkey MD, Weiser SD, Fehmie D, Alamo-Talisuna S, Sunday P, Nannyunja J, et al. Socioeconomic determinants of mortality in HIV: evidence from a clinical cohort in Uganda. J Acquir Immune Defic Syndr 2014; 66. 41–47.

39. Delpierre C, Cuzin L, Lauwers-Cances V, Datta GD, Berkman L, Lang T. Unemployment as a risk factor for AIDS and death for HIV-infected patients in the era of highly active antiretroviral therapy. Sex Transm Infect 2008; 84: 183–6.
40. DeSilva MB, Merry SP, Fischer PR, Rohrer JE, Isichei CO, Cha SS. Youth, unemployment, and male gender predict mortality in AIDS patients started on HAART in Nigeria. AIDS Care 2009; 21: 70–7.
41. Sartorius B, Kahn K, Collinson MA, Sartorius K, Tollman S. Dying in their prime: determinants and space-time risk of adult mortality in rural South Africa. Geospat Health 2013; 7: 237–49.
42. MacPherson P, Moshabela M, Martinson N, Pronyk P. Mortality and loss to follow-up among HAART initiators in rural South Africa. Trans Roy Soc Trop Med Hyg 2009; 103: 588–93.

Citation: Glob Health Action 2014, **7**: 25488 - http://dx.doi.org/10.3402/gha.v7.25488

Understanding perceived risk factors for mortality from friends and relatives of demised HIV-infected people in Soweto, South Africa: A qualitative study

Kennedy N. Otwombe[1,2], Max Petzold[2,3], Neil Martinson[1,4], Tobias Chirwa[2], Janan J. Dietrich[1]

Affiliations

[1]Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[3]Health Metrics, University of Gothenburg, Gothenburg, Sweden

[4]Johns Hopkins University, Center for TB Research, Baltimore, MD, USA

Author's Email Addresses

Kennedy N. Otwombe: otwombek@phru.co.za

Abstract

Background: The burden of HIV remains high in South Africa where up to 7 million people are living with HIV. Of these, just over half are on treatment yet the benefits of treatment in relation to HIV associated mortality are well known. Although HIV associated mortality has lowered, the number still dying is significant. Hence this qualitative study explores perceived risk factors associated with HIV mortality in which persons that are living, talk about their demised friends or relatives.

Method: Twenty participants undertook in-depth interviews between May and August 2014 at a clinic in the Perinatal HIV Research Unit.

Results: The 20 enrolled participants ages ranged from 18 to 63 years; 17 were females. Participants reported knowing the deceased persons for between 5 and 40 years. Perceived risk factors for mortality included risky sexual behaviour, negative attitude by healthcare workers, religious prayers and traditional medicine, food security and social support structure.

Conclusion: Our findings showed the importance of providing social support to HIV-infected individuals. Social support such as helping in domestic chores, food security and support to move around was perceived importantly as a factor that may have reduced the burden of disease and potentially alleviate the risk of mortality. More research is required to show the association of social support and mortality in HIV infected people.

Keywords: HIV; Perceived risk factors; Mortality;

Introduction

HIV remains the greatest burden of disease in South Africa where approximately 7 million people live with HIV (PLWH) and more than half are currently on treatment.(1) Access to treatment with good adherence counselling services has been critical in reducing HIV associated mortality. While the number of PLWH increased between 2005 and 2010, mortality decreased by more than half within the same period.(1) The decrease in mortality is associated with provision of free antiretroviral therapy (ART) in public health care facilities amongst other reasons.

Risk factors associated with mortality in HIV infected individuals have been widely researched and include socio-economic as well as clinical factors. Under socio-economic factors, gender, employment and age have been reported; clinical factors have included low CD4 count, high viral load, tuberculosis and other opportunistic infections have previously been associated with HIV mortality.(2-7) Additionally, substance use and non-adherence to treatment have also been reported. Treatment guidelines have been updated to recommend immediate initiation of HIV treatment following a positive test but it is not unusual to find several HIV infected individuals initiating therapy with poor immunological and virological outcomes.(5, 6)

Understanding of perceived risk factors associated with HIV mortality amongst friends and relatives of demised HIV infected people has not been well explored. Often, research on HIV related perceptions is conducted on a sample of people to understand their perception of various aspects associated with HIV such as stigma, social norms, HIV testing and health care workers amongst others.(8-10) Research

is limited on key informant interviews of relatives and friends of HIV infected people that demised seeking their perceptions on potential risk factors for mortality.

There is a paucity of qualitative research focussing on perceptions of risk factors for HIV mortality in which persons that are living, talk about their demised friends or relatives. Our research is unique because the reported perceptions focus on areas that may have elevated the risk of death in the demised HIV infected individuals. Additionally, perception was assessed in relation to level of illness, home based care, financial support, food security, treatment and treatment adherence and health care workers.

Qualitative Methods

Study Setting and Population

The present study was conducted at the Perinatal HIV Research Unit (PHRU), a division of the University of the Witwatersrand and based at the Chris Hani Baragwanath Hospital in Soweto. Soweto is an urban area in the city of Johannesburg in Gauteng, South Africa with a population above one million people.(11) Soweto is in the South western part of Johannesburg and approximately 40% of Johannesburg's population reside there.(12) While many households are poor with high unemployment rates, there is a mix of poor and wealthy residents.(12) The overall HIV prevalence of the province where Soweto is located is 12.4% and females are disproportionately affected.(13) While mortality in HIV infected people has declined as a result of free antiretroviral treatment, deaths due to late initiation of treatment and tuberculosis still remain a concern.(2, 3, 14)

Sampling

Participants were recruited through the PHRU ZAZI voluntary counselling and testing (VCT) centre. ZAZI is a successful HIV testing service that is offered for free to clients in Soweto and surrounding areas. In the present study, participants were purposively recruited from the ZAZI VCT. Of the 26 participants invited to participate, 20 were enrolled. During the VCT counselling process, clients were routinely asked about the loss of a person in their life and probed further to determine if it was due to HIV. Those responding to the affirmative were probed further by counsellors to determine who it was, when it happened and how they dealt with the loss. They were also asked if the deceased person was HIV infected and if they were, are invited to participate in this study. Those willing to participate provided their telephone contact and thereafter interviews were scheduled within a period of seven days. A day prior to the scheduled interview date, a telephone call was made to confirm the appointment. At the start of the interview, we asked how close participants were to the deceased persons and how often they saw each other. Most of them were related with relations such as aunt, uncle, brother, sister or cousins. We also asked whether the deceased persons were in a relationship such as marriage prior to death.

Eligibility

To be eligible, participants had to be 18 years or older, have known a close person who had died due to HIV and were willing to participate in a face-to face in-depth interview.

Ethical considerations

Study procedures were approved by the Human Research and Ethics Committee of the University of the Witwatersrand. We obtained written informed consent to participate in the study from all eligible participants. Consenting occurred at two levels: for participation in the study and audio-recording. The consent forms and interview guides were written in English but the consenting process and the in-depth interviews were carried out in either English or Zulu as appropriate. Participants were re-imbursed approximately $15 (~ 150 ZAR) for their transport costs. Participants were informed that counselling will be provided by counsellors based at the ZAZI VCT should they be distressed during the interview.

In-depth interviews

Data were collected between May and August 2014 using in-depth interviews to understand the characteristics of the participant and their deceased loved one. A demographic questionnaire assessed age, gender, primary home language and educational level and was administered prior to the in-depth interview.

Interviews were conducted in English and/or isiZulu by a trained interviewer who had experience in conducting qualitative interviews. The in-depth interviews explored participants perceptions about: the deceased persons' relationships with significant others or in general, health seeking behaviours, attitude of healthcare workers towards HIV-infected patients seeking care, initiation and adherence to antiretroviral therapy (ART), home-based care support, nutrition and/or food security and financial stability. Interviews occurred in private rooms at the ZAZI VCT centre and lasted between 45 and 60 minutes.

All interviews were audio-recorded and transcribed verbatim one day after the interview. Transcriptions were performed by an external team with experience in

transcribing audio recordings that was approved by the study team. Where audio-recordings had a mixture of the local language and English, translations into English occurred during the transcription process. Every second transcript was selected by the interviewer for comparison with the audio recording to check for consistency. Where discrepancies were identified, audio-recordings were transcribed again. A total of 10 transcripts were selected for comparison with the audio-recordings for quality checks (that were done by a bilingual research assistant).

All the study procedures were recorded in a study work book that was used to monitor study progress. It contained information such as how many people had been interviewed, how many more were likely to be interviewed, scheduled interview appointments and administrative information such as participant re-imbursements for transport.

Data analyses

Data were analysed through a framework analysis approach where similar and interrelated codes were grouped together to form sub-themes and themes in a data matrix.(15, 16) Core codes were independently developed by two research assistants and agreed upon in collaboration with an experienced qualitative researcher. The first four transcripts were reviewed by two research assistants line by line and the findings used to develop a code book. A meeting was organised between the research assistants and a qualitative analysis expert to verify the consistency of the codes. Divergent codes were discussed and a resolution reached on the most appropriate coding term. Subsequent transcripts were assessed based on the code-book. Updated codes were only included if subsequent codes could not fit within the initial ones.

The data matrix was instrumental in assessing the coding progress allowing for the development of sub-themes and themes.(15, 16) The process of developing sub-themes and themes was regularly refined during the analytic process as new codes emerged. Any discrepancies in coding were resolved through consensus. Development of codes, sub-themes and themes followed a combined approach of deductive and inductive data reduction procedures.

Trustworthiness of the data

Trustworthiness of the data was assessed using various methodologies.(17-20) An audit trail was kept to monitor the study progress such as evaluating the first two interviews and amending the interview guides as appropriate. Codes were independently prepared by the research assistants and harmonised in a meeting together with an experienced qualitative researcher. Thereafter, a codebook was developed for all subsequent interviews. Code development stopped when no further changes occurred. Audio-recordings were compared with the transcripts to verify consistency, ensure translations matched the recordings and make any appropriate amendments. Clarity during the interview process was achieved through probing of responses by the research assistant. Final results were shared with the study team for verification of the themes and interpretation. Differences in the coding, themes and interpretations were resolved as a team.

Results

Demographics

Of the 20 participants that were enrolled, their minimum and maximum ages were 18 and 63 years respectively and 17 were females. Their highest school grade was between 2nd and 12th Grades and 9 reported Isizulu as their main language. They

reported living in households with between 1 and 9 people and the total rooms in the house ranging from 1 to 17. There were 10 and 7 participants from female headed households where the head of household was aged 18-60 and > 60 years respectively. Participants reported knowing the deceased persons for between 5 and 40 years. Eight participants were employed, 8 were unemployed and 4 were students in tertiary institutions. Deceased persons were either an aunt (n=5), cousin (n=7), son (n=2) or other relation (n=6).

Participants reported that the deceased persons were aged between 18 and 54 years, 8 reported IsiZulu as their main language and their highest school grade was between 2 and 12. Of the 20 deceased persons, 15 were females. Among the deceased persons, 12 were employed, 6 were unemployed and 2 were students in tertiary institutions.

Emerging themes on perceptions of mortality risk factors

Following the interviews, several themes emerged that were perceived as having an association with mortality among the deceased persons. Participants perceived that engagement in risky sexual behaviour such as multiple sexual partnerships may have a role in worsening the situation for a HIV-infected person. A a negative attitude by healthcare workers towards HIV-infected people was perceived to be associated with avoidance in seeking care. Believing in the healing power of religion while ignoring antiretroviral treatment and traditional medicine were perceived to have worsened the conditions of the deceased persons. Food security such as availability of regular meals as well as a good social support structure such as receiving help in domestic chores was perceived as important in keeping the HIV-infected persons healthy.

Perceived engagement in risky sexual behaviour

Participants reported about the deceased's risky behaviour such as multiple partnerships, unprotected sex without condoms and alcohol use. Some of the deceased were perceived to be involved in multiple relationships whilst being married. For instance, a participant stated the following of her father:

> *"I stayed with him for that 6 years and he was alright but you could see that even when he was here in Joburg, he still had a gift of liking women…he used to take me with him when he went to see his women (girlfriends). He was aware that he is with me. He had 4 (wives)." (Female Participant 1 in relation to her deceased father)*

Another provided the following summary of her friend's husband:

> *"She told me about (her husband) from what he does and so on. She said (her husband) would come back home in the morning and he knew very well that we (husband and deceased) are sick but he still did the things he did. That was reason why I moved out of the house cause he made me sick. He should have done the right thing. I joked about it and told her that she shouldn't have had just one boyfriend in life till they got married. She (deceased) said that is one thing that hurt her and made her sad. On the other hand, her partner (husband of the deceased) went on and slept with multiple partners" (Female participant 6 in relation to her deceased friend's husband)*

Participants commonly reported that once the deceased persons had improved while on treatment, they defaulted on their treatment. Thereafter they went back to their

previous habit of drinking alcohol and engaging in multiple partnerships. A participant stated the following of her aunt:

> *"Yes she (deceased person) recovered then she stopped taking the treatment and went back to drinking again, continued with her boyfriends and that's where she fell ill and got worse." (Female participant 12 in relation to her deceased aunt)*

Participants reported that engagement in unprotected sex occurred and it was unclear if disclosure of their HIV status happened. A female participant quoted the following of her aunt:

> "*Now she (deceased person) is in a relationship with a pastor who has a wife and they don't use a condom and who knows if (my aunt) told him that she is sick*" *(Female participant 6 in relation to her deceased aunt)*

It was perceived that the deceased persons engaged in multiple partnerships and unprotected sex while knowing their HIV status. Furthermore, it was unclear if they disclosed their HIV status to their partners.

Negative attitude by healthcare workers

A negative attitude by the healthcare workers was perceived to be a barrier in seeking care in public health facilities. The quote below summarises an advise given by a clinician to the deceased aunt of a participant:

"He (the doctor) told her to terminate (her pregnancy) but she didnt want to and she never went back to the doctor and that is why she never went to the clinic to get medication to protect the baby from getting infected." (Female participant 6 in relation to her aunt)

Another participant reported on the perceived negative attitude shown to her deceased relative by the nurses in the local clinic. The perceived negative attitude to a sister-in-law is summarised below:

"She (deceased person) said that nurses were not treating her well. I asked her why she didn't tell us because we could ask our aunt to take her to the (local) clinic and you (the deceased) would take the treatment there. She said that she was afraid. I asked her, how long would she be afraid when she had kids who were already complaining that their mother is always sleeping and they didn't understand what was going on." (Female participant 16 in relation to her sister-inlaw)

Sometimes when participants reported on the perceived negative attitude by clinicians, they failed to be specific. A female participant provided the following summary of her deceased aunt's experience:

"*The only thing she (deceased person) told me was that the doctors are rude. That is the only thing I knew. People usually used to say that. But she (deceased person) would never explain to you why she's saying they are rude. I would ask why you (deceased person) say they are rude and she*

would say they are rude. And the nurses when you call them, they take their time (to respond)" (Female participant 9 in relation to her deceased aunt"

It was perceived that the negative attitude by healthcare workers discouraged the deceased persons from seeking care in the public health care facilities and that may have led to deteriorating health.

Preference for religion and traditional medicine

The use of religious prayers as well as traditional medicine for purposes of healing was espoused during the interviews. Others perceived withcraft as the cause of their HIV infection. They believed someone bewitched them into their status.

The perceived use of religious prayers as a measure of healing by an aunt is summarised in the following quote by a female participant:

"She (deceased) believed in the church. She never went to sangomas (traditional healer). She went (to the church) because she loved it. She believed in it. She knew even if she was HIV (infected), God is there. She believed that miracles do happen" (Female Participant 6 in relation to her deceased aunt)

Others attended churches where the clergy provided holy water or tea. Holy water refers to water that has been blessed by the clergy and apparently contains healing powers. Similarly holy tea is blessed tea. This is summarised in the following quotes by relatives of deceased persons:

"She (deceased person) believed in church. They asked the pastor and the congregation to come to our house. Then they prayed for her (deceased person). And they prayed for the (holy) water and told her how she must drink it and she must wash with it." (Female Participant 10 in relation to her deceased cousin)

"He had a belief that since he is a member of (name of church) that he will get healed (by drinking tea). It was tea that they got from the church. How can you get healed by drinking tea? I even asked him (deceased person). Has he ever seen anyone get healed by drinking tea" (Male Participant 11 in relation to his deceased brother)

The use of traditional medicine in managing HIV emerged during the interview. It was perceived that traditional medicine, provided by sangomas (traditional healers), was favoured by either the deceased person or their families. This is encapsulated by relatives of deceased persons in the following quotes:

"You know they even tried to consult a Sangoma to help her to get better. You know the traditional healers said we need to slaughter a chicken and make a ceremony for her. Oh and the Sangoma said she must drink the chicken's blood." (Female Participant 9 in relation to her deceased aunt)

"They (family) did take him to a traditional healer and he was given traditional tonic and medication and things like that. I am not sure, but they took him to a

traditional healer because of the chest pains suspecting that he was bewitched in his beer. You know how people are? So they suspected witchcraft because he was complaining of pain at the side of his chest." (Female Participant 20 in relation to her deceased aunt)

Other participants reported that some family members perceived withcraft as the cause of the illnesses suffered by the deceased. One participant summarised the following of her aunt:

"No I don't think she (deceased person) told them because they did not visit her while she was sick. My aunt suspected witchcraft and because I knew the situation, I told her not to point fingers because I was here when mother was very sick. I saw her. Some people are like that. They blame witches for everything. When things go wrong, when their children fail at school, they say it is witchcraft." (Female Participant 8 in relation to her deceased mother)

The significance attached to religious prayers, traditional medicine and witchcraft emerged during the interview. These were perceived as avenues that offered alternative forms of care for HIV infected individuals.

Food security

During the interviews, it emerged that food security was perceived as a matter of concern for HIV infected individuals that had to stop working due to HIV associated opportunistic infections. A female participant summarised her mother's experience as follows:

"When she started coughing seriously, she had to quit (working). It was her decision. Her body became weak and she was not coping. She was very sick and feeling pain she could not go to work." (Female Participant 8 in relation to her deceased mother)

Unfortunately not working led to loss of income followed by reduced availability of food especially for those who were self-employed and running their own businesses. A participant stated the following of her deceased aunt:

"She used to sell food in a shop (in Soweto). So when she got sick, she couldn't go sell. So there was no one running her business. She was used to doing things on her own. That actually stressed her because on the other hand she didnt have money. So my uncle that lives in Pimville (township in Soweto) and other aunt that lives with us helped her. They told her if she needs anything, they will help. They told her not to worry but to focus on her health. That is when she relaxed a bit." (Female Participant 7 in relation to her deceased aunt)

Participants reported that some of the deceased persons were unemployed and relied on others such as parents, friends or partners for food. A female participant summarises the experience of her deceased cousin as follows:

"Her mother would bring her food but then she said there were times when she had nothing to eat. So she would have to go eat with her friends or her

boyfriends, something like that." (Female Participant 2 in relation to her deceased cousin)

Not only did the deceased persons stop working and impacting on household food security but also those who cared for them. As one participant noted, of the mother of one of the deceased persons:

"Yes, then she (mother of deceased) stopped (working) when her daughter got sick.Yes, she (mother of deceased) stopped working so that she can take care of…(her daughter) and her child (grandchild)" (Female Participant 5 in relation to her deceased cousin)

The perceived importance of food security and its association with income emerged during the interview. Those who stopped working due to illness lost their income. Participants said that the support provided by the family and friends enabled the deceased persons an opportunity to live without worrying about food.

Social support structure

The perceived importance of the social support structure such as the family, relatives and friends emerged during the interviews. Discussions revealed that deceased persons were able to receive home-based care in the form of support in attending to domestic chores or attending clinics. While other deceased persons received family support, others did not. A female participant noted the following of her aunt who never got support from the boyfriend:

"The only thing that I heard was about her (aunty's) boyfriend. She (aunty) complained that since she started being sick, he (boyfriend) didn't come and

see her. He wasn't there for her". (Female Participant 7 in relation to her deceased aunt)

There were those who became weak, frail and bedridden and required support with domestic cores such as washing their dirty linen. As one female participant noted of her aunt, the frail required help.

"They did everything for her. Her father would wake up early in the morning to make porridge for her and her mother would give her a bath. She would change her linen for her. She was very clean. Every Sunday, they would get someone to drive her to church because she liked church. They would drive her everywhere she would want to go " (Female Participant 6 in relation to her deceased aunt)

Participants reported that support to move around and take treatment on time was possible due to the availability of a social support structure.

"She made food for her. Made sure that she takes her pills on time. If she needed her to help her stand, she would do that for her. The pills that she was taking was grandpas (analgesics) that she bought herself at the spaza shop (informal store). She would boil water for her and also rub her feet if they needed to get rubbed until she went to the hospital" (Female Participant 7 in relation to her deceased aunt)

The perceived importance of a social support structure to provide support to frail HIV-infected persons was emphasised during the interviews. The social support structure enabled the deceased HIV-infected persons lead a comfortable life.

Additionally taking treatment timely was assured due to the presence of a person to ensure that medication is taken.

Discussion

Our study was conducted to determine the perceived factors associated with mortality in deceased HIV-infected people with a close relationship to the participants enrolled in the study. Emerging themes included risky sexual behaviour, negative attitude by healthcare workers, religious prayers and traditional medicine, food security and social support structure.

It was perceived that the risk of mortality in the deceased persons may have been elevated by engagement in risky sexual behaviour such as multiple sexual partnerships and unprotected sex. The role of multiple sexual partnerships as well as unprotected sex in relation to HIV and new HIV cases has widely been researched.(21-26) Additionally, evidence is now available to show that HIV infection is likely to occur in stable relationships.(22, 25) Unfortunately partners only become aware of their HIV status when their health condition deteriorates. Often this leads to late diagnosis and initiation of therapy when the immune system is severely immunosuppressed leading to early mortality.(5, 6) While antiretroviral therapy is provided for free in public health facilities, mortality rates associated with HIV still remain high.

Negative attitude by healthcare workers was perceived to have discouraged the deceased persons from seeking care in the public healthcare facilities. This finding is consistent with existing literature about discrimination of HIV-infected people by healthcare workers.(8, 27-29) While there has been more awareness about management of HIV among healthcare workers,(9) some still display a negative attitude towards HIV-infected individuals. Unfortunately discriminatory practice against HIV-infected individuals dissuades them from seeking care in public

healthcare facilities. This in turn worsens their health condition and may lead to death as a result of improper care and management. Additionally, discrimination of HIV-infected individuals leads to poor quality of life due to low levels of treatment adherence and low utilisation of healthcare services.(8, 9, 30)

In our findings, participants perceived that some deceased person had more faith in healing through religious prayers. They believed that going to church, asking for forgiveness and praying might lead to a miracle that would reverse their HIV condition. Indeed, this was reinforced by religious leaders willing to offer prayers and support towards healing. Furthermore, some clergy even presented a form of blessed water suggesting that consuming it would lead to healing from HIV. The use of religion as a coping mechanism for HIV has previously been documented.(31, 32) While religion frames the way of life for many people including HIV-infected individuals, the choice of ignoring care and initiation of therapy for those eligible to start treatment for the sake of religious prayers may be detrimental to their health. It may worsen the health of immunosuppressed individuals leading to negative consequences such as death.(5, 6) It may be that healthcare providers need to consider an integrated approach that provides care to HIV-infected individuals while at the same time accommodating their religious way of life such as allowing them to continue seeking healing through prayers. The interplay between religion and the belief that it leads to healing of HIV requires further research.

We found preference for traditional medicine over antiretroviral therapy by some of the deceased persons. The perceived benefits associated with the use of traditional medicine in treating HIV remains unclear. However it may be that its use is associated with the need to identify with one's indigenous identity, family expectations, privacy and confidentiality.(33, 34) Furthermore use of traditional

medicine in combination with antiretroviral therapy may lead to drug interactions(35) and poor health outcomes such as death. Healthcare providers should consider use of both traditional and contemporary medicine under their supervision in order to mitigate any adverse events and manage them accordingly.(33) This should be feasible within systems where regulations governing use of traditional medicine are in place such as in South Africa.

The combined effect of a HIV-related death and food insecurity has previously been studied where it has been shown that social and development implications cannot be ignored.(36, 37) Participants perceived a threat to food security in the livelihood of deceased persons due to loss of income when they ceased employment. This was often due to HIV-associated illnesses. Mortality is disproportionately high in the HIV-infected people who are often economically active.(36, 38, 39) As a result, the livelihood of their dependents is put at risk. Our findings concur with previously reported research showing the association between food insecurity and mortality in HIV-infected individuals.(37, 40-42)

Limitations in this study may include social desirability bias where participants were more likely to say things that portray them positively to the interviewer. All of the participants were VCT clients and thus likely to follow a line of thought introduced to them by the VCT counsellors. For some of the participants, it is likely that recall bias may have affected their responses since the demised HIV infected person may have died a few years earlier.

Our findings showed the importance of providing social support to HIV-infected individuals. Participants perceived that provision of social support such as helping in domestic chores, food security and support to move around may have reduced the burden of disease and potentially reduced the risk of mortality. While literature on

social support such as home-based care for HIV-infected people is widely available,(4, 10, 43-45) there is limited information if any to show how provision of social support is associated with mortality in HIV-infected individuals.(46) More research is required to determine the effect of social support structure on morbidity and mortality of HIV-infected people especially in sub-Saharan Africa where the burden of disease is high.

Acknowledgement

We are grateful to all the participants that enrolled in this study and allowed their data to be used for research purpose and the counsellors at the clinical for their role in conducting in-depth interviewst.

Financial support

This research was supported by the Perinatal HIV Research Unit and the Consortium for Advanced Research Training in Africa (CARTA). CARTA is jointly led by the African Population and Health Research Center and the University of the Witwatersrand and funded by the Wellcome Trust (UK) (Grant No: 087547/Z/08/Z), the Department for International Development (DfID) under the Development Partnerships in Higher Education (DelPHE), the Carnegie Corporation of New York (Grant No: B 8606), the Ford Foundation (Grant No: 1100--0399), Google.Org (Grant No: 191994), Sida (Grant No: 54100029) and MacArthur Foundation Grant No: 10-95915-000-INP.

KO received an academic scholarship from Canadian-Africa Prevention Trials Network and the Forgaty International Center (grant no. TW007370/3).

References

1. UNAIDS. UNAIDS data. Accessed on 23rd March 2018 at http://www.unaids.org/sites/default/files/media_asset/20170720_Data_book_2017_en.pdf. 2017.

2. MacPherson P, Moshabela M, Martinson N, Pronyk P. Mortality and loss to follow-up among HAART initiators in rural South Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2009;103(6):588-93.

3. Martinson N, Variava E. Mortality in TB patients and active case finding in HAART treated adults. Accessed on 31 July 2014 at file:///C:/Users/otwombek/Downloads/TC_S___Dr_Neil_Martinson___Mortality_in_TB_patients_and_active_case_finding_in_HAART_treated_adults.pdf 2008 [

4. Massavon W, Barlow-Mosha L, Mugenyi L, McFarland W, Gray G, Lundin R, et al. Factors Determining Survival and Retention among HIV-Infected Children and Adolescents in a Community Home-Based Care and a Facility-Based Family-Centred Approach in Kampala, Uganda: A Cohort Study. ISRN AIDS. 2014;2014:852489.

5. Mujugira A, Wester CW, Kim S, Bussmann S, Gaolathe T. Patients with advanced HIV type 1 infection initiating antiretroviral therapy in Botswana: treatment response and mortality. AIDS research and human retroviruses. 2009;25(2):127-33.

6. Otwombe KN, Laher F, Tutu-Gxashe T, Gray G, Mohapi L. The effect of a maturing antiretroviral program on early mortality for patients with advanced immune-suppression in Soweto, South Africa. PLoS One. 2013;8(11):e81538.

7. Otwombe KN, Petzold M, Modisenyane T, Martinson NA, Chirwa T. Factors associated with mortality in HIV-infected people in rural and urban South Africa. Glob Health Action. 2014;7:25488.

8. Famoroti TO, Fernandes L, Chima SC. Stigmatization of people living with HIV/AIDS by healthcare workers at a tertiary hospital in KwaZulu-Natal, South Africa: a cross-sectional descriptive study. BMC medical ethics. 2013;14 Suppl 1:S6.

9. Holzemer WL, Makoae LN, Greeff M, Dlamini PS, Kohi TW, Chirwa ML, et al. Measuring HIV stigma for PLHAs and nurses over time in five African countries. SAHARA J : journal of Social Aspects of HIV/AIDS Research Alliance / SAHARA , Human Sciences Research Council. 2009;6(2):76-82.

10. Young SD, Hlavka Z, Modiba P, Gray G, Van Rooyen H, Richter L, et al. HIV-related stigma, social norms, and HIV testing in Soweto and Vulindlela, South Africa: National Institutes of Mental Health Project Accept (HPTN 043). Journal of acquired immune deficiency syndromes. 2010;55(5):620-4.

11. City of Johannesburg. Soweto. Accessed on 7th April 2015 at http://census2011.adrianfrith.com/place/798026 2011 [

12. Wikipedia. Soweto. Accessed on 7th April 2015 at http://en.wikipedia.org/wiki/Soweto 2015 [

13. Shisana O, Simbayi LC, Zuma K, Jooste S, Zungu N, Labadarios D, et al. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Accessed on 18 April 2015 at http://www.hsrc.ac.za/uploads/pageContent/4565/SABSSM%20IV%20LEO%20final.pdf. 2014.

14. Wubshet M, Berhane Y, Worku A, Kebede Y, Diro E. High loss to followup and early mortality create substantial reduction in patient retention at antiretroviral treatment program in north-west ethiopia. ISRN AIDS. 2012;2012:721720.

15. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology. 2013;13.

16. Smith JM, Firth J. Qualitative data analysis: the framework approach. Nurse Researcher. 2011;18(2):52-62.

17. Loh J. Inquiry into issues of trustworthiness and quality in narrative studies: a perspective. The Qualitaive Report 2013;18(65):1-15.

18. Morrow SL. Quality and trust worthiness in qualitative research in counseling psychology. Journal of Counseling Psychology. 2005;52(2):250-60.

19. Carcary M. The research audit trail-enhancing trust worthiness in qualitative inquiry. The Electronic Journal of Business Research Methods. 2009;7(1):11-24.

20. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Education for Information. 2004;22:63-75.

21. Bellan SE, Fiorella KJ, Melesse DY, Getz WM, Williams BG, Dushoff J. Extra-couple HIV transmission in sub-Saharan Africa: a mathematical modelling study of survey data. Lancet. 2013;381(9877):1561-9.

22. Exavery A, Kante AM, Tani K, Hingora A, Phillips JF. Sociodemographic drivers of multiple sexual partnerships among women in three rural districts of Tanzania. Hiv/Aids. 2015;7:105-13.

23. Scorgie F, Chersich MF, Ntaganira I, Gerbase A, Lule F, Lo YR. Socio-demographic characteristics and behavioral risk factors of female sex workers in sub-saharan Africa: a systematic review. AIDS and behavior. 2012;16(4):920-33.

24. Serwadda D, Wawer MJ, Musgrave SD, Sewankambo NK, Kaplan JE, Gray RH. HIV risk factors in three geographic strata of rural Rakai District, Uganda. Aids. 1992;6(9):983-9.

25. Shisana O, Zungu-Dirwayi N, Toefy Y, Simbayi LC, Malik S, Zuma K. Marital status and risk of HIV infection in South Africa. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2004;94(7):537-43.

26. WHO. The World Health Report 2002. Reducing Risks, Promoting Healthy Life. Accessed on 17 June 2015 at http://www.who.int/whr/2002/en/whr02_en.pdf. 2002.

27. Ahsan Ullah AK. HIV/AIDS-Related Stigma and Discrimination: A Study of Health Care Providers in Bangladesh. Journal of the International Association of Physicians in AIDS Care. 2011;10(2):97-104.

28. Horsman JM, Sheeran P. Health care workers and HIV/AIDS: a critical review of the literature. Social science & medicine. 1995;41(11):1535-67.

29. Knussen C, Niven CA. HIV AIDS and health care workers: Contact with patients and attitudes towards them. Psychol Health. 1999;14(3):367-78.

30. Letamo G. The discriminatory attitudes of health workers against people living with HIV. PLoS medicine. 2005;2(8):e261.

31. Cotton S, Puchalski CM, Sherman SN, Mrus JM, Peterman AH, Feinberg J, et al. Spirituality and religion in patients with HIV/AIDS. Journal of general internal medicine. 2006;21 Suppl 5:S5-13.

32. Trevino KM, Pargament KI, Cotton S, Leonard AC, Hahn J, Caprini-Faigin CA, et al. Religious coping and physiological, psychological, social, and spiritual outcomes in patients with HIV/AIDS: cross-sectional and longitudinal findings. AIDS and behavior. 2010;14(2):379-89.

33. Belisle HA, Hennink M, Ordonez CE, John S, Ngubane-Joye E, Hampton J, et al. Concurrent use of traditional medicine and ART: Perspectives of patients,

providers and traditional healers in Durban, South Africa. Glob Public Health. 2015;10(1):71-87.

34. Puoane TR, Hughes GD, Uwimana J, Johnson Q, Folk WR. Why HIV positive patients on antiretroviral treatment and/or cotrimoxazole prophylaxis use traditional medicine: perceptions of health workers, traditional healers and patients: a study in two provinces of South Africa. African journal of traditional, complementary, and alternative medicines : AJTCAM / African Networks on Ethnomedicines. 2012;9(4):495-502.

35. Babb DA, Pemba L, Seatlanyane P, Charalambous S, Churchyard GJ, Grant AD. Use of traditional medicine by HIV-infected individuals in South Africa in the era of antiretroviral therapy. Psychology, health & medicine. 2007;12(3):314-20.

36. Lemke S. Nutrition security, livelihoods and HIV/AIDS: implications for research among farm worker households in South Africa. Public health nutrition. 2005;8(7):844-52.

37. Twine W, Hunter LM. Adult mortality and household food security in rural South Africa: Does AIDS represent a unique mortality shock? Development Southern Africa. 2011;28(4):431-44.

38. Gregson S, Mushati P, Nyamukapa C. Adult mortality and erosion of household viability in AIDS-afflicted towns, estates, and villages in eastern Zimbabwe. Journal of acquired immune deficiency syndromes. 2007;44(2):188-95.

39. Sartorius B, Kahn K, Collinson MA, Sartorius K, Tollman SM. Dying in their prime: determinants and space-time risk of adult mortality in rural South Africa. Geospatial health. 2013;7(2):237-49.

40. Dowdy DW, Geng EH, Christopoulos KA, Kahn JS, Hare CB, Wlodarczyk D, et al. Mortality Among Antiretroviral-Eligible Patients in an Urban Public Clinic. Jaids-J Acq Imm Def. 2011;57(4):297-300.

41. Kalichman SC, Cherry C, Amaral C, White D, Kalichman MO, Pope H, et al. Health and Treatment Implications of Food Insufficiency among People Living with HIV/AIDS, Atlanta, Georgia. J Urban Health. 2010;87(4):631-41.

42. Weiser SD, Fernandes KA, Brandson EK, Lima VD, Anema A, Bangsberg DR, et al. The Association Between Food Insecurity and Mortality Among HIV-Infected Individuals on HAART. Jaids-J Acq Imm Def. 2009;52(3):342-9.

43. Kabore I, Bloem J, Etheredge G, Obiero W, Wanless S, Doykos P, et al. The effect of community-based support services on clinical efficacy and health-related quality of life in HIV/AIDS patients in resource-limited settings in sub-Saharan Africa. AIDS patient care and STDs. 2010;24(9):581-94.

44. Knight LC, Van Rooyen H, Humphries H, Barnabas RV, Celum C. Empowering patients to link to care and treatment: qualitative findings about the role of a home-based HIV counselling, testing and linkage intervention in South Africa. AIDS care. 2015:1-6.

45. Root R, Whiteside A. A qualitative study of community home-based care and antiretroviral adherence in Swaziland. Journal of the International AIDS Society. 2013;16:17978.

46. Mohanan P, Kamath A. Family support for reducing morbidity and mortality in people with HIV/AIDS. The Cochrane database of systematic reviews. 2009(3):CD006046.

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M130818

NAME: (Principal Investigator) Mr Kennedy Otwombe

DEPARTMENT: School of Public Health
Medical School

PROJECT TITLE: Determing Predicators of Mortality in HIV Positive People in South Africa, 2003 to 2009: A Mixed Methods Approach Incoporating Hidden Variables

DATE CONSIDERED: 30/08/2013

DECISION: Approved unconditionally

CONDITIONS:

SUPERVISOR: Prof Tobias Chirwa

APPROVED BY:

Professor PE Cleaton-Jones, Chairperson, HREC (Medical)

DATE OF APPROVAL: 12/03/2014

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

yearly progress report.

Principal Investigator Signature M130818Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Document Viewer

Turnitin Originality Report

Processed on: 31-Mar-2018 9:06 PM SAST
ID: 939097202
Word Count: 26956
Submitted: 1

0508147a:Kennedy-Thesis_Draft_31MAR2018_KO_Fi... By Kennedy Otwombe

1% match (Internet from 19-May-2016)
http://www.globalhealthaction.net

1% match (publications)
Otwombe, Kennedy N., Max Petzold, Neil Martinson, and Tobias Chirwa. "A Review of the Study Designs and Statistical Methods Used in the Determination of Predictors of All-Cause Mortality in HIV-Infected Cohorts: 2002–2011", PLoS ONE, 2014.

Similarity Index	Similarity by Source	
11%	Internet Sources:	9%
	Publications:	6%
	Student Papers:	2%

1% match (publications)
Chang, Sherilyn, Yunjue Zhang, Anitha Jeyagurunathan, Ying Wen Lau, Vathsala Sagayadevan, Siow Ann Chong, and Mythily Subramaniam. "Providing care to relatives with mental illness: reactions and distress among primary informal caregivers", BMC Psychiatry, 2016.

<1% match (Internet from 23-Sep-2015)
http://www.globalhealthaction.net

<1% match (Internet from 24-Jul-2016)
http://journals.plos.org

<1% match (student papers from 21-Sep-2016)
Submitted to University of Witwatersrand on 2016-09-21

<1% match (Internet from 01-May-2008)
http://www.foodsafetymatters.gov.au

<1% match (publications)
Courtenay Sprague. "Gender and HIV in South Africa", Springer Nature, 2018

<1% match (Internet from 16-Nov-2016)
http://journals.lww.com

<1% match (Internet from 27-Aug-2016)
http://mobile.wiredspace.wits.ac.za

<1% match (Internet from 06-Sep-2017)
http://paduaresearch.cab.unipd.it

<1% match (Internet from 28-Dec-2017)
http://clok.uclan.ac.uk

<1% match (Internet from 30-May-2016)
http://acikerisim.deu.edu.tr

<1% match (Internet from 27-May-2016)
http://repub.eur.nl

<1% match (publications)
Wienke, . "Univariate Frailty Models", Chapman & Hall/CRC Biostatistics Series, 2010.

<1% match (Internet from 24-Oct-2017)
http://circoutcomes.ahajournals.org

<1% match (Internet from 12-Sep-2017)
http://digitalcommons.wustl.edu

<1% match (publications)
"Alternatives for the frailty model", The Frailty Model, 2008