61

Survival analysis of under-five twins in Ethiopia: a gamma frailty modeling approach

By: Mesfin Tsegaye

A Thesis submitted in partial fulfillment for the requirements of

Master of Science (MSc) Degree in Biostatistics

In the Department of Statistics, Faculty of Natural Sciences,

Addis Ababa University

June 2013

Advisor: - Dr. Emmanuel Gabreyohannes

Addis Ababa University
School of Graduate Studies

This is to certify that the thesis prepared by Mesfin Tsegaye Haileyesus, entitled: Survival analysis of under-five twins in Ethiopia: a gamma frailty modeling approach and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Statistics (Bio-Statistics) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Examiner Mekonnen Tadesse Signature _____ Date 27/06/2013

Examiner _____ Signature _____ Date 27/06/2013

Advisor _____ Signature _____ Date _____

_____

Chair of Department or Graduate Program Coordinator

Abstract

Twins are relatively rare events, but several studies confirm that they contribute substantially to mortality in both neonatal and post-neonatal periods. The excess infant and child mortality rates among twins calls for a need to identify the main causes behind it. This study intended to identify risk factors that are significantly associated with survival of under-five twins in Ethiopia. Data about twin under-five mortality was found from the birth history of women who were included in the 2011 Ethiopia Demographic and Health survey. This study, therefore, employed bivariate survival analysis approach using gamma frailty models. The results of the study showed that place of residence, preceding birth interval, birth order, and previous child status were significantly associated with under-five twin mortality.

Acknowledgement

I want to thank my advisor Dr. Emmanuel Gabreyohannes for his valuable advice and support throughout the preparation of this thesis. This thesis could not have been written without his unreserved help and support. I am also tremendously grateful to Mr. Belay Birele[1], who opened my eyes for survival analysis and guided me throughout this thesis. I wish to thank Dr. Iachine I.A for providing me with some important papers for this thesis. I want to thank my mother and Ms. Senayit Seyoum[2] for the unreserved financial support and encouragement. Last, but not least, I want to thank my family and all of my friends for their support and encouragement.

[1] Mr. Belay Birele: Instructor in Jimma University

[2] Ms. Senayit Seyoum: Instructor in Haromaya University

Table of contents

List of Tables

List of Figures

Chapter one

1. Introduction

1.1Background of the study

Despite the significant improvements in child survival in the past decades, levels of infant and child mortality and morbidity remain high in many developing countries (UNICEF, 2010). These problems are particularly serious among high-risk pregnancies and births, and in developing countries where the health-care system is still struggling to provide basic public health and maternal and child health care to their population (Bryce et al, 2005).

Twins are relatively rare events, but contribute substantially to mortality in both neonatal and post-neonatal periods (Alam et al, 2007). There are few studies that evaluate the mortality of twins after the neonatal period in Sub-Saharan African countries. Those studies include the studies brought together by Pison (1992), Justesen and Kunst (2000), Olalekan et al (2008), Aaby et al (1995). All of these studies examine the excess mortality of twin compared to singleton.

The exact risk factors behind under age 5 twins mortality are not yet clearly known, but it is possible to think that either genetic or environmental related factor plays vital role in survival of individual twin. Children belonging to the same family share certain unobserved characteristics, which may not be sufficiently described by the observed covariates included in statistical models. The ignorance of such family-level correlation may lead to biased parameters estimates. This unobserved heterogeneity, also referred to as frailty (Vaupel et al, 1979), operates at three different levels: child, family and community (Sastry, 1997). At the family level, children from the same parents inherit common genetic factors and usually grow up in the same household environment. Parents are also more likely to adopt similar child care behavior for all their children. Genetic factors remain the major component of the family-level frailty. However, each child has a proper susceptibility to infection, independently of his family membership (Childs et al, 1992). This idiosyncratic genetic factor remains the major child level unobserved frailty component. In addition, inside the common global family behavioural factor, parents may adopt a slightly different prenatal and neonatal attitude from one child to the next, for example, the

health care practice and the nutritional status. At the community level, the random effects are more likely of behavioural and environmental nature (Koissi and Högnäs, 2001).

The persistence of high infant and child mortality rates among twins calls for a need to identify the main causes behind this phenomenon. Twin and adoption studies have provided insight on the relative importance of genes and environment on the variation of a trait. Twins represent a very special case of relationship. There are two types of twins: Monozygotic (MZ) and Dizygotic (DZ) twins. MZ twins are genetically identical, whereas DZ twins share 50 % of their genes on average like ordinary siblings. This difference can help to evaluate the relative importance of genes and environment on the variation of a trait.

1.2 Statement of the problem

According to The Ethiopian Gemini Trust (2010) Ethiopia has twice as many twin births as Europe. However, without assistance and problems related to nutrition almost 30 % of the babies die before they turn one year old. Furthermore, the same source reports that in Ethiopia the infant mortality rate among twins is at least three times higher than singleton.

The rate of occurrence of infant and child mortality rates reflects the country's level of socio economic development and quality of life. Although child mortality rate shows encouraging decline pattern in the past 10 years, if Ethiopia come across to achieve the MDG 4 (reduce child mortality by two-thirds between 1990 and 2015) child mortality rate must be further reduced. Therefore, in addition to in-depth understanding of the levels, trends, differentials and determinants of child mortality, identification of determinants or the risk factors particularly for twin under-5 mortality is also vital in any attempt to attain the goal through proper and sustainable types of intervention.

To the best of our knowledge, no study has assessed the risk factors that aggravate twin under-5 mortality in Ethiopia, although there are several studies that have been conducted on infant and child mortality such as Asefa et al (2000), Deribew et al (2007), and Fottrell et al (2009) among others. Most of these studies have used Cox proportional hazard model to estimate the significance of the risk factors on child mortality. Some of these studies have also employed indirect estimation of Brass and Trussell's methods: It estimates mortality with an assumption that the risk of child death is a function of the age of the child only and does not depend on other

factors such as mother's age or child's birth order. These studies assumed independence between observations and ignored the correlation between children from the same family. However, in studies involving multiple individuals from the same family, the independence assumption is not plausible unless all important familial factors were measured and controlled for in the model. Furthermore, they failed to include the effect of unmeasured variability, and hence, some unmeasured genetic, environmental and components remain unexplained.

1.3 Objectives of the study

The general objective of the study is to identify risk factors that are significantly associated with survival of under-five twins in Ethiopia.

- **Specific objectives**
  - To fit a parametric correlated gamma frailty model with observed covariates.
  - To assess if there are unobserved environmental and genetic factors that aggravate the twin under-five mortality in Ethiopia.

1.4 Significance of the study

Evidence-based action plans and interventions are needed to reduce child mortality. Recently, maternal, environmental, behavioral and socioeconomic factors were recognized as important factors that aggravate child mortality. The result of this study could be used as knowledge input for any attempt to reduce child mortality through proper and sustainable types of interventions.

1.5 Limitations of the study

- Among the available covariates in the study dataset (The 2011 Ethiopia Demographic and Health survey), those information (taken at the time of interview) on father/mother's education level, and wealth index had been omitted because they might have changed during the time preceding the interview (time varying variable). Neither of mother's marital status, the household variables such as family size, the source of drinking water and the type of toilet used for the same reason.

- The study dataset has no information on the zygosity of the twins, which could have able us to assess the unobserved factor effects in detail. Thus, in this study the cause of the unobserved factor is not assessed.
- Since it was not possible to get the appropriate statistical software to fit bivariate parametric proportional hazards models it was not possible to include those tasks that require high level programs such as model assessment in the parametric section.

Chapter two

2. Literature review

Traditionally bivariate survival data have been studied by analysis-of-variance methods developed by quantitative geneticists Bulmer (1980) and Falconer (1990). A new approach has been opened up by the extension of the methods of survival analysis by Kalbfleish and Prentice (1980) and Cox and Oakes (1984) to bivariate survival data. The survival analysis approach is superior to the analysis-of-variance approach when the data to be analyzed are censored and advantageous when the mechanisms of mortality, including the effect of covariates, can be appropriately captured by a hazards model (Yashin et al, 1995).

Recently, investigators have recognized that ignoring individual heterogeneity may lead to inaccurate conclusions. Models for heterogeneity have been proposed by Vaupel et al (1979), who introduced frailty as an unobserved quantity in population mortality. Frailty refers to a susceptibility to death that is not captured by observed covariates. Typically, frailty includes factors that affect an individual's survival chances such as genes and unmeasured attributes of the environment, all of which may or may not be shared to some degree with other individuals in a family. The most common frailty specification is the ''shared frailty'' extension of the proportional hazards regression model. Hougaard et.al (1992) used several versions of the frailty model for bivariate survival to fit data for Danish monozygotic and dizygotic twins. In a shared frailty model, the frailties are unobserved random variables assumed to be independent and to follow a probability distribution, the shape of which is described by few parameters. The shared frailty model pools all effects, shared genetic and shared environmental, into a single random effect without consideration of the genetic relationships that link any two relatives.

Oakes (1989) proposed frailty models for bivariate survival times and introduced several possible frailty models. Yashin and Iachine (1995) and Yashin et.al (1995) introduced a model of bivariate survival that allows for the incorporation of correlations between individual frailties. These models are known as ''correlated frailty'' models, which have been used with twin data to assess genetic and environmental factors influencing mortality.

Salihu et al (2005) examined the association between maternal nativity and neonatal survival of twins among black mothers in USA. They compared the fetal growth patterns of twins born to US and foreign born black mothers by computing means of gestational age-specific birth weights for each group and plotting these against their respective gestational ages. Then they assessed overall differences between the two groups by means of the Student *t* test. Comparisons of levels of overall, early, and late neonatal mortality between the two groups are made using hazard ratios generated from a Cox Proportional Hazards Regression model. After testing for the non-violation of the proportionality assumption, adjusted and unadjusted hazard ratios were generated by using the partial likelihood method described by Cox (1972). They also employed the Robust Sandwich Estimator (Liang and Zeger, 1986) to adjust the estimates of the variance of the coefficients in order to account for the correlation among observations within twin sets. Their study determined the morbidity pathway for adjusted differences in neonatal mortality between the two nativity groups. The three precursors of neonatal mortality, low birth weight, preterm, and small-for-gestational-age, were tested as candidate mediators that could have accounted for any observed differences in neonatal demise between the two ethnic sub-populations. All tests of hypotheses were two-tailed, with a type I error rate fixed at 5%. The result of the study shows that twins of US born mothers had a 23% higher likelihood of dying within the neonatal period compared to those of foreign-born mothers (hazard ratio = 1.23, 95% confidence interval: 1.04 - 1.46). The disparity in neonatal demise occurred exclusively in the early neonatal period (hazard ratio = 1.29, 95% CI: 1.06 - 1.50), with mortality indices comparable in the late neonatal period (hazard ratio = 0.96, 95% CI: 0.68 - 1.35). Low and very low birth weight ($p<.0001$), preterm and very preterm ($p<.0001$), and small-for-gestational-age neonates ($p<.0001$) were more prevalent among twins of US-born mothers. Finally they concluded that, compared to those of foreign born, twins of US born black mothers experienced higher mortality in the neonatal period. The mortality disadvantage resulted mainly from lower gestational age at birth and the preponderance of small-for-gestational-age babies among US-born black mothers.

Garibotti et al (2006) studied longevity and correlated frailty in multigenerational families. The study applied Cox proportional hazards models to data from three-generation pedigrees in the Utah Population Database using two different frailty specification schemes that account for

common environments (shared frailty) and genetic effects (correlated frailty). In a model that includes measures of familial history of longevity and both frailty effects, they found that the variance component due to genetic factors is comparable to the one attributable to shared environments: Standard deviations of the correlated and the shared frailty distributions are 0.143 and 0.186, respectively. Through simulations, they also showed a greater reduction in the bias of parameter estimates for fixed covariates through the use of the correlated frailty model.

The application of the bivariate correlated gamma frailty model with observed covariates to the lifetimes of Monozygotic (MZ) and Dizygotic (DZ) Danish twins with respect to coronary heart disease (CHD) were performed by Wienke and others and expressed in Wienke (2011). They analyzed the influence of smoking, body mass index (BMI), gender, and birth year (using the variable transformation birth year minus 1890 because the oldest twins on this study were born in 1890) on the susceptibility to death by CHD. They fitted parametric models with different cumulative baseline hazards such as Weibull, Exponential and Gompertz. Based on the likelihood, compared to the other two, the Gompertz model showed the best fit to the data. The correlation for MZ twins ($\hat{\rho}_{Mz}$= 1) was on the boundary of the parameter space [0, 1], whereas DZ twins showed a smaller correlation ($\hat{\rho}_{Dz}$ = 0.689). Cigarette smokers had a significantly higher risk to die from CHD, with $\hat{\beta}$= 0.465 and standard error SE=0.149, compared to nonsmokers. Twins with small or high BMI have a slightly worse prognosis compared to twins in the reference category with BMI between 22 and 28 kg/m2. Women showed a significantly better survival with $\hat{\beta}$ = −0.783 (SE=0.126) compared to males, and later birth is related to a decreasing risk.

A semiparametric bivariate correlated gamma frailty model was fit to MZ and DZ Danish twins born between 1890 and 1920 by Pietzner and Wienke (2010). Here the aim was to analyze the influence of smoking, body mass index (BMI), gender, and birth year on mortality. The correlation for MZ twins was $\hat{\rho}_{Mz}$ = 0.706 (SE=0.086), whereas DZ twins showed a smaller correlation $\hat{\rho}_{Dz}$ = 0.479 (SE=0.130).

Wienke et al (2003) fit a correlated gamma-frailty model to study the genetic influence on susceptibility to diseases of the respiratory system and all-cause mortality. This study used data from the Danish twin registry. Twin pairs born between 1870 and 1930, where both individuals

were alive on 1 January 1943 were included in this study. The data was categorized by sex and zygosity information. And proportions of variance in frailty attributable to genetic and environmental factors were assessed using structural equations model (Neale and Cardon, 1992). The result of the study showed that, for all-cause mortality the correlation coefficients of frailty for MZ twins tend to be higher than for DZ twins. And for mortality with respect to respiratory diseases this effect was only seen in females, whereas males showed the opposite effect. In addition to this, after they fit the five standard biometric models (Neale and Cardon, 1992) to the data, the analysis confirms the presence of a strong genetic influence on individual frailty associated with all-cause mortality. But for respiratory diseases, no genetic influence was found in males and only weak genetic influence in females.

Roudsari et al (2006) compared the risk of early childhood injuries in twins and singletons. They conducted a retrospective cohort study using linked birth certificate, hospital discharge, and death certificate data from Washington State (1987–2002). They used Cox proportional hazards regression to determine the hazard ratio (HR) of injury-related hospitalization and death in the first six years of life for twins and triplets compared to singletons. Then the HRs of injury-related hospitalization were 1.4 for twins (95% CI: 1.2–1.6) relative to singletons after adjustment for factors such as child's sex, mother's age, marital status, and number of older siblings. Finally, they showed that low birth weight significantly modified the association between twin status and injury hospitalization, and they conclude that the risk of injury death was not significantly higher among twins than singleton children. Rather twins appear to be at higher risk of childhood injury hospitalization.

Becher et al (2004) used Cox proportional hazard model to quantify childhood mortality in rural Burkina Faso. This study used data from demographic surveillance system in 39 villages around Nouna, Western Burkina Faso. All children born alive in the period 1 January 1993 to 31 December 1999 were included. The result of the study shows that death of mother and being a twin are the strongest risk factors for child mortality.

Fottrell et al (2009) used Demographic and Health Survey (DHS) and Demographic Surveillance System (DSS) data from Ethiopia to model the distribution and effects of under-five mortality risk factors. They employed Cox proportional hazard model to derive the hazard rate ratio for the

two data sources. The results of the study show that child mortality risk profile was similar between each data source, with multiple births and living in less populous households are significant risk factors for under-five mortality.

Desta (2011) used data from 2000 and 2005 Ethiopian DHS and employed logistic regression analysis to examine the socioeconomic, demographic and biological factors of infant and child mortality. The study show that marital status, birth order, type of births and preceding birth intervals are significant proximate determinants of infant and child mortality. Breast feeding had an important significant effect on infant mortality but not on child mortality. Children born to women not married, first born children, multiple birth, children born within 18 months of the previous birth and children who were breastfed for less than 6 months were exposed to a higher risk of infant and child mortality. Children born in small household size, children born in male headed household, children born to mothers and fathers with no education and to some extent children born to mothers and fathers with primary education were exposed to a higher risk of infant and child mortality.

Joshua and Jeroen (2009) used 2005-06 Zimbabwean DHS to investigate the maternal, socioeconomic and sanitation factors on infant and child mortality using Cox regression model. The result of the study shows that birth order six and more with short preceding interval is significantly associated with higher risk of infant and child mortality. Multiple births increase infant and child mortality. Children who are first born and those born to mothers aged 40-49 years were associated with higher infant and child mortality. They showed that the influence of birth order, preceding birth intervals, maternal age, type of birth and sanitation factors have a pronounced effect on infant mortality but weak effect on child mortality.

Wang (2003) using the 2000 Ethiopia DHS examines the environmental determinants of child mortality by constructing three hazard models (the Weibull, the Piece-wise Weibull and the Cox model). This study examine three age-specific mortality rates: neonatal, infant, and under-five mortality by location (rural/rural), female education attainment, religion affiliation, income quintile, and access to basic environmental services (water, sanitation and electricity). The estimation results show a strong statistical association between child mortality rates and poor environmental conditions.

Balk et al (2003) used DHS data from 12 West Africa countries and carry out a spatial analysis on childhood mortality using logistic regression model. The result of the study showed that children who are first born were associated with higher risk of death during infancy compared to children who have birth order above 5. Multiple births were associated with higher risk of death during infancy. Children born to educated mother (secondary or higher education) experience high survival. Furthermore, infants and children who reside in urban areas showed a better survival chances than those children reside in rural areas.

Njagi and Purity (2011) used data from the 2008/09 Kenya Demographic Health Survey to examine the implications of unobserved heterogeneity on parameter estimates of urban rural differentials in infant mortality. In this study Standard Log normal Accelerated Failure Time model was used for analysis, and shared frailty model was further fit to account for unobserved heterogeneity. Frailty was considered at household level. The result of the study indicates that in urban areas birth size was the only significant factor that influenced infant mortality, whereas in rural areas region of residence, preceding birth interval, birth order and birth size were significant factors.

Niragire et al (2006) fit Cox proportional hazard model with shared frailty models to identify the determinants of child mortality in Rwanda. They used data from the 2005 Rwanda Demographic Health Survey. The result of the study shows that frailty effects were significant in childhood; with child deaths mostly determined by socioeconomic and demographic factors such as household socioeconomic status.

Sastry (1997) apply a multivariate proportional hazards model with nested frailty to analyze the effect of covariates on child survival. This study used survey data from northeast Brazil collected via a hierarchically clustered sampling scheme. The study shows that family and community frailty effects were fairly small in magnitude but important. And children born from women at youngest and oldest age are subject to high risk of death.

Guo and Rodríguez (1992) assess the effects of covariates and analyze the effects of unobserved family heterogeneity in children survival times. They fit a shared gamma frailty model with observed covariates. The study dataset was obtained from a retrospective survey conducted on

1974-76 in Guatemala. The result of the study showed that children born from women at youngest age were at highest risk of death and short proceeding and succeeding birth intervals increase child mortality risk. In addition to this, the result of the study confirmed the significance of family random effect.

Koissi and Högnäs (2001) analyze the effects of unobserved family heterogeneity in children survival times. The analysis was done through a Bayesian approach using a proportional hazard model with multiplicative random effect. They used data from the Demographic and Health Survey that was carried out from September 1998 to March 1999 in Ivory Coast. The result of the study showed that children whose previous sibling died experienced a lower survival chance. When the reference group has a preceding birth interval greater than 24 months, a preceding birth interval less than 18 months increases the child mortality risk by 2 times. First births were also 1.5 times riskier. The family random effect was found statistically significant.

Chapter three

3. Data and Methodology

3.1 Data and data source

The source of the data used in this study was the 2011 Ethiopia Demographic and Health Survey (EDHS). It was conducted in Ethiopia as part of the worldwide demographic and health survey project. The 2011 Ethiopia Demographic and Health Survey was conducted by the Central Statistical Agency (CSA) with the support of the Ministry of Health. This is the third Demographic and Health Survey (DHS) conducted in Ethiopia under the worldwide measure DHS project, a USAID-funded project providing support and technical assistance in the implementation of population and health surveys in countries worldwide.

DHSs are large, complex cross-sectional surveys that measure demographic and health parameters on a nationally representative sample. Nationally distributed cluster samples of households are performed at approximately five-year intervals, with each round drawing a new cross-section sample. The standard DHS survey consists of a household questionnaire and a women's questionnaire administered to a nationally representative sample of women aged 15-49 years. The women's questionnaire is used to gather information on complete birth histories to estimate infant and child mortality probabilities.

The 2011 Ethiopia Demographic and Health survey interviewed a nationally representative population in about 18,500 households, and all women of age 15-49 and all men of age 15-59 in these households. Indicators relating to family planning, fertility levels and determinants, fertility preferences, infant, child, adult and maternal mortality, maternal and child health, nutrition, women's empowerment, and knowledge of HIV/AIDS are provided for the nine regional states and two city administrations. In addition, data by urban and rural residence at the country level are provided.

Data about twin mortality was found from the birth history of women who were included in the survey. All twin births before January 01, 2008 were included.

3.2 Variables of the study

- **The Response variable**

The response or outcome variable for this study is the survival time of a pair of under-five twins measured in weeks.

- **Independent variables**

Among the various covariates that are thought to have effect on survival of a child the factors birth order, gender, mothers age at birth, preceding birth interval, previous child survival status and residence are included in this study. For the purpose of interpretation those quantitative covariates are changed into categorical covariates. The categories of each covariate and the corresponding reference categories are presented in Table 3.1.

Table 3. 1: The covariates and the categories of the covariates in the study

Covariates	Categories	Code
Residence	Urban Rural	1 2
Sex	Male Female	1 2
Age of mother at birth	Below 18 years Between 18- 35 years Above 35 years	1 2 3
Preceding birth interval	Below 18 month Between 18-24 month Above 24 month or first born	1 2 3
Birth order	First born Between 2 -4 5 and above	1 2 3
Previous child status	Dead Alive or 1st child	1 2

3.3 Survival analysis

Survival analysis is a statistical method for data analysis where the outcome variable of interest is the time to the occurrence of an event. Time to an event is a positive real valued variable having continuous distribution. It is necessary to define the starting time point; the occurrence of an event may be death, occurrence of disease, time to an epileptic seizure, time it takes for a patient to respond to a therapy, or time from response until disease relapse. It is applied in a number of applied fields, such as medicine, public health, social science, and engineering. In medical science, time to event can be time until recurrence in a cancer study, time to death, or time until infection. In the social sciences, interest can lie in analyzing time to events such as job changes, marriage, birth of children and so forth.

Survival time refers to the time from a particular starting point to a particular end point of interest or occurrence of event. The survival time data set contain either censored or truncated observations.

Censored data arise when an individual's life length is known to occur only in a certain period of time. Well-known censoring schemes are right censoring, where all that is known is that the individual is still alive at a given time; left censoring, when all that is known is that the individual has experienced the event of interest prior to the start of the study; or interval censoring, where the only information is that the event occurs within some interval.

Well-known truncation schemes are left truncation, where only individuals who survive a sufficient time are included in the sample and right truncation, where only individuals who have experienced the event by a specified time are included in the sample.

In addition survival time data is usually skewed and always non-negative, which implies that the normality assumption is unlikely to be satisfied. Such data cannot be handled properly by standard statistical methods. Researchers use different techniques to respond to the complication due to censoring but until recently none of the techniques was entirely satisfactory. However, new developments in statistical theory accompanied by new development in statistical computing have changed how researchers can study such data. This new method known as survival analysis was developed by biostatisticians modeling human life times (Cox, 1972, Cox

and Oaks 1984; Kalbfleish and Prentice, 1980; Miller, 1981) in the medical and biological sciences.

3.3.1 Survival Function

Let T be a random variable denoting the survival time. The distribution of survival times is characterized by any of three functions: the survival function, the probability density function or the hazard function. The following discussion is due to Hanagal (2010).

Survival function $S(t)$ is defined as:

$S(t) = P[T > t]$ = the probability an individual survives beyond time t.

Since a unit either fails or survives, and one of these two mutually exclusive alternatives must occur, due to the above definition of $S(t)$ we have:

$$S(t) = 1 - F(t) \qquad (3.1)$$

where $F(t)$ is the cumulative distribution function (CDF) of T. If T is a continuous random variable, then $S(t)$ is a continuous, strictly decreasing function. The survival function is the integral of the probability density function $f(t)$ of T, that is:

$$S(t) = \int_t^{\infty} f(x)dx \qquad (3.2)$$

Thus $$f(t) = -\frac{dS(t)}{dt} \qquad (3.3)$$

3.3.2 Failure (or Hazard) Rate

The hazard function gives the instantaneous failure rate at t given that the individual has survived up to time t, i.e.

$$h(t) = lim_{h\to 0} \frac{P[t \le T \le t + h|\mathrm{T} \ge \mathrm{t}]}{h}$$

$$= \frac{f(t)}{S(t)} = instantaneous(conditional)\ failure\ rate$$

The failure rate is sometimes called a "conditional failure rate" since the denominator $S(t)$ (i.e., the population survivors) converts the expression into a conditional rate, given survival past time t. Since $h(t)$ can be expressed as:

$$h(t) = \frac{d}{dt} ln\{S(t)\}$$

it follows that:

$$S(t) = exp\left\{-\int_0^t h(t)dt\right\} \quad (3.4)$$

If we define the cumulative Hazard Function by:

$$H(t) = \int_0^t h(t)dt$$

then, we have

$$S(t) = e^{-H(t)} \quad (3.5)$$

There are many general shapes for the hazard rate, the only restriction on $h(t)$ is that it be non negative. The hazard rate for the occurrence of a particular event can be increasing, decreasing, constant, bathtub-shaped, hump-shaped, or possessing some other characteristic which describe the failure mechanism.

The bivariate survival function of the lifetimes (T_1, T_2) is given by:

$$S(t_1, t_2) = P[T_1 > t_1, T_2 > t_2] = exp[-H(t_1, t_2)], \quad (3.6)$$

where $H(t_1, t_2)$ is the bivariate integrated hazard function of (T_1, T_2) which can be written in terms of bivariate survival function as:-

$$H(t_1, t_2) = -lnS(t_1, t_2) \quad (3.7)$$

3.3.3 Approaches in estimation of Survival function

To estimate the survival function, two different modeling approaches are used, namely parametric and non parametric. In the case of parametric estimation, it is necessary to make assumptions about the distribution of failure times. In some circumstances this makes sense, especially when additional information about the nature of the underlying aging or disease process is available. Some of the important models widely used in various research papers are the exponential, Weibull, gamma, log normal, log logistic, normal, exponential power and Gompertiz distribution. Each parametric distribution is defined by a different hazard function.

In the non parametric case, no assumption is required about the distribution of failure times. Kaplan Meier estimator (Kaplan and Meier, 1958), Nelson-Allen estimator (Nelson, 1969) and Berslow estimator (Breslow, 1974) are the most widely used non parametric methods.

Whether to use a parametric or a nonparametric model is an important point. An advantage of nonparametric models is their flexibility and the resulting ability to deal with any probability distribution. However, there is a high price to pay. First, nonparametric methods need much more data to get reasonable results. Second, it is hard to get estimates of the hazard function, which is often an interesting and relevant information. In contrast, parametric models often allow closed-form expressions of the hazard and survival function depending on the chosen model. Parametric models can be described by the values of a few parameters. They often give good results even in the case of small sample size. If the assumed model is correct, the estimation is more efficient than in a nonparametric estimation procedure.

As a general approach one should plot the hazard function for the observed data and determine whether or not it is consistent with the assumed parametric distribution. If the data follows a parametric distribution, parametric methods are preferred to non-parametric methods for describing and quantifying factors that influence time to event.

3.4 Cox Proportional hazards model

The proportional hazards model introduced by Cox (1972) is a regression model with event time as the dependent variable. It allows the inclusion of information about known (observed) covariates in models of survival data in an easy way and is the most applied model in this area.

Let $h(t \mid X)$ denote the hazard of an individual at time t with covariate $X' = (X_1, X_2, \dots, X_k)$. The proportional hazards model specifies that:

$$h(t \mid X) = h_0(t) G(X) \qquad (3.8)$$

where $h_0(t)$ is the baseline hazard function and $G(.)$ some positive function.

The model assumes a baseline hazard that all individuals in the study population have in common. The parameters of primary interest are contained in $G(X) = G(\beta, X)$, with $\beta' = (\beta_1, \beta_2, \dots, \beta_k)$ denoting the vector of regression parameters. In this model, the covariates act

multiplicatively on the baseline hazard. This allows the model a simple and easy interpretation. It is assumed that all individual variation in the hazard can be characterized by a finite-dimensional vector of observed covariates.

The main idea behind proportional hazards model is the separation of the time effect in the baseline hazard function on the one hand and the effect of the covariates in an exponential term on the other. In essence, this assumption says that the hazards of two individuals at time t is related by a proportionality constant that does not depend on t. The simple two-sample situation is obtained by restricting to a single (k = 1) binary covariate X in the model with X = 0 or X = 1, depending on group membership. In this case, the method is truly nonparametric, and e^{β} denotes the hazard ratio between the two groups. However, if X is continuous, a parametric form of $G(.)$ is required. Inference is now dependent on that parametric form but still independent of $h_0(t)$, and the model is called a semiparametric model because of the parametric nature of the covariate term and the nonparametric baseline hazard function. In the semiparametric Cox ph model the survival function given the covariates X is given by:

$$S(t \mid X) = S_0(t)^{e^{\beta \cdot X}} = e^{-\mathrm{H_0(t)}e^{\beta \cdot X}} \qquad (3.9)$$

where $S_0(t) = e^{-\mathrm{H_0(t)}}$ denotes the baseline survival function, $\mathrm{H_0(t)}$ denotes the baseline cumulative hazard function, and the components of the vector β are unknown regression parameters. That means the survival function of an individual with covariate vector X is a power of the baseline survival function. The class of distributions generated by this procedure is sometimes called Lehmann alternatives (Lehmann, 1953).

To estimate β and $\mathrm{H_0(t)}$ from data that are censored, Cox (1972, 1975) proposed a method called partial likelihood method. The partial likelihood expression in (3.10) does not depend on $h_0(t)$. Suppose that, data are available for n individuals, among whom there are r distinct death times and $n - r$ right-censored survival times. Assume that only one individual dies at each death time, so that there are no ties in the data. Let $t_{(1)}<t_{(2)}< \ldots <t_{(r)}$ denote the r ordered death times. To estimate the regression coefficients β, the partial likelihood function is given by:

$$l(\beta) = \prod_{j=1}^{r} \frac{\exp(\beta' X_{(j)})}{\sum_{l \in Rt_{(j)}} \exp(\beta' X_l)} \qquad (3.10)$$

where $X_{(j)}$ is the vector of covariates for the individual who dies at the j^{th} ordered death time $t_{(j)}$. The summation in the denominator of the function (3.10) extends over all individuals who are at risk at time $t_{(j)}$. Based on the estimates of β, one can obtain an estimate of $H_0(t)$ by using the Berslow baseline cumulative hazard estimator (Breslow, 1974). The details about the estimation procedure with tied events and the likelihood construction is available in Lee and Wang (2003).

3.4.1 Test of proportional hazards assumption

The main assumption of the Cox proportional hazards model is proportional hazards. Proportional hazards means that the hazard function of one individual is proportional to the hazard function of another individual, i.e., the hazard ratio is constant over time. There are several methods to verify as to whether a model satisfies the assumption of proportionality.

As a procedure one can run a Cox model with each covariate (individually) and introduce a time-dependent interaction term for that covariate. If the proportional hazards assumption is valid for the covariate, the time-dependent interaction term should not be significant. This approach is regarded as the most sensitive (and objective) method for testing the proportional hazards assumption (Stevenson, 2009). In the same way, we can assess the PH assumption for several predictors simultaneously.

3.4.2 Cox proportional hazards model diagnostics

In linear regression methods, residuals are defined as the difference between the observed and predicted values of the dependent variable. However, when censored observations are present and the partial likelihood function is used in the Cox PH model, the usual concept of residual is not applicable. A number of residuals have been proposed for use in connection with the Cox PH model. We will describe three residuals in the Cox model.

1. **Cox-Snell residuals**

The Cox-Snell residual (Cox and Snell, 1968) for the i^{th} individual with observed survival time $t_{(i)}$ is defined as:

$$r_{ci} = \exp(\hat{\beta}' X_i) * \widehat{H_0}(t_{(i)}) = -\log\left(\hat{S}(t_{(i)})\right) \qquad (3.11)$$

where $\hat{\beta}$ is the partial likelihood estimator of the regression coefficient and $\widehat{H_0}(t_i)$ is an estimate of the Breslow's baseline cumulative hazard function Breslow (1974) at time $t_{(i)}$ which is given by:

$$\widehat{H_0}(t_i) = \sum_{i:t_i<t} \frac{\delta_i}{\sum_{j\epsilon R(t_i)} e^{\beta'(k)X_j}}$$

This residual is motivated by the following result: Let T has a continuous survival distribution $S(t)$ with the cumulative hazard $H(t) = -\log(S(t))$. Thus, $S_T(t) = \exp(-H(t))$. Let $Y = H(t)$ be a transformation of T based on the cumulative hazard function. Then the survival function for Y is given by:

$$\begin{aligned} S_Y(Y) &= P(Y > y) = p(H(t) > y) \\ &= P\left(T > H_T^{-1}(y)\right) = S_T\left(H_T^{-1}(y)\right) \\ &= \exp\left(-H_T\left(H_T^{-1}(y)\right)\right) = \exp(-y) \end{aligned}$$

Thus, regardless of the distribution of T, the new variable $Y = H(t)$ has an exponential distribution with unit mean. Therefore, we use a plot of $H(r_{ci})$ versus r_{ci} to check the fit of the model. This gives a straight line with unit slope and zero intercept if the fitted model is correct. Note that the Cox-Snell residuals will not be symmetrically distributed about zero and cannot be negative.

2. Martingale Residuals

Martingale residuals are the difference between the observed number of events for an individual and the conditionally expected number given the fitted model, follow up time, and the observed course of any time-varying covariates. Martingale residuals may be plotted against covariates to detect non-linearity (that is, an incorrectly specified functional form in the parametric part of the model). Martingale residuals (sometimes referred to as Cox-Snell or modified Cox-Snell residuals) are defined as:

$$M_i = \delta_i - r_{ci} \qquad (3.12)$$

where δ_i is censoring indicator defined as $\delta_i = 1$ if the observed survival time t_i is uncensored and 0 otherwise. If the plot of Martingale residuals versus the covariate do not show any pattern, the linearity assumption for the covariate under consideration is satisfied.

3. **Deviance residuals**

The deviance residuals help us in identifying poorly fitted subjects, and is defined (Therneau et al., 1990) as:

$$D_i = sign(M_i)\sqrt{-2[M_i + \delta_i \log(\delta_i - M_i)]} \qquad (3.13)$$

where sign (.) is the sign function which takes the value 1 if M_i is positive and -1 if M_i is negative. The deviance residuals are symmetrically distributed about zero when the fitted model is adequate, and individuals with large positive or negative deviance residuals are poorly predicted by the model.

Dfbeta

Dfbeta is a useful measure to assess the influence of each observation on the estimated coefficients $\hat{\beta}$s. This measure is analogous to that used in the usual linear regression. Large values suggest we inspect the corresponding data points.

3.5 Parametric proportional hazard model

Parametric proportional hazards model is the parametric version of the Cox proportional hazards model. It has a similar form to that of Cox PH models. The main difference between the two kinds of models is that the baseline hazard function is assumed to follow a specific distribution when a fully parametric PH model is fitted to the data, whereas the Cox model has no such constraint. The coefficients are estimated by partial likelihood in Cox model whereas the maximum likelihood approach is used in parametric PH model. Other than this, the two types of models are equivalent. Hazard ratios have the same interpretation and proportionality of hazards is still assumed.

Here we discuss parameter estimation in the bivariate parametric PH models. The details about the univariate parametric PH model parameter estimation are available in Jiezhi Qi (2009). Let us consider the simple case-when the paired observations are independent. The bivariate survival function of the lifetimes (T_1, T_2) is given by:

$$S(t_1, t_2) = S_1(t_1)S_2(t_2) \qquad (3.14)$$

Assume identical survival function for the two individuals in a pair i.e $S_1(t_1) = S_2(t_2) = S(t)$. And suppose Weibull cumulative baseline hazard $H_0(t) = \lambda t^v$ (where $\lambda>0$ and $v > 0$ are the scale and shape parameters, respectively). Then the bivariate survival function of the lifetimes (T_1, T_2) with observed covariate vectors X_1 and X_2 is given by:

$$S(t_1, t_2|X_1, X_2) = \exp\left(-\lambda {t_1}^v \times e^{\beta \cdot X_1}\right) \exp\left(-\lambda {t_2}^v \times e^{\beta \cdot X_2}\right) \qquad (3.15)$$

Then, as expressed in Kalbfleish and Prentice (2002), the complete likelihood function is given by:

$$\prod_{i=1}^{n} S(dt_1, dt_2)^{\delta_{i1}\delta_{i2}} - S(dt_1, t_2)^{\delta_{i1}(1-\delta_{i2})} - S(t_1, dt_2)^{(1-\delta_{i1})\delta_{i2}} S(t_1, t_2)^{(1-\delta_{i1})(1-\delta_{i2})} \qquad (3.16)$$

where δ_{ij} denotes the censoring information ($\delta_{ij} = 0,1$; $\delta_{ij} = 0$ indicates right censoring) for the j^{th} individual in the i^{th} pair($j = 1,2; i = 1,2, \ldots, n$) and dt denotes the derivative with respect to t. Using standard maximum likelihood procedure, one can maximize the log likelihood and obtain the estimates of the parameters in the model. One can follow a similar procedure for Gompertz and exponential baseline hazard functions.

3.6 Concept of frailty

Sometimes it is difficult to include all covariates that have impact on the survival time of an individual. There are two main reasons why it is often impossible to include all important factors on the individual level into the analysis. Sometimes there are too many covariates to be considered in the model, in other cases the researcher may not know or is not able to measure all the relevant covariates. But these unobserved variable(s) may have significant effect on the dependent variable (survival time) (Hanagal, 2010).

The random effect, or frailty models, have been introduced into the statistical literature in an attempt to account for the existence of unmeasured attributes such as genotype that do introduce heterogeneity into a study population.

The term frailty was introduced by Vaupel et al. (1979) to indicates that different individuals are at risks even though on the surface they may appear to be quite similar with respect to measurable attributes such as age, gender, weight, etc. Frailty may be common to all study group (shared) or different from individual to individual.

Frailties are useful in modeling correlations in multivariate survival and event history data. Frailties can be nested (individuals within a family may share a common frailty, while families within communities share another common frailty) and it can be correlated.

A common approach to the analysis of survival data is to assume a homogeneous population of individuals with the same covariate structure. However, it is clear that individuals identical in many respects such as age, sex, and treatment may differ in unmeasured ways, if only because of genotypical differences. Ignoring such heterogeneity has a negative consequence in parameter estimation. For details see Hanagal (2010).

Frailty distribution describes the frailty in the population at the start of the follow-up. Frailty is assumed to be fixed for each individual over time, but the composition of the population changes as time goes by. On average, more frail individuals die earlier. Due to this fact the frailty distribution in the population at risk changes over time (Wienke, 2011).

The classical and most frequently applied model of frailty assumes a proportional hazards structure that is conditional on the random effect (frailty). To be more specific, the hazard function of an individual depends on an unobservable, time-independent random variable Z called frailty which acts multiplicatively on the baseline hazard function $h_0(t)$ (Wienke, 2006). The conditional hazard function given Z has the form:

$$h(t \mid z) = Zh_0(t) \tag{3.17}$$

Here, Z is considered as a nonnegative random mixture variable, varying across the population. Note that a scale factor common to all subjects in the study population may be absorbed into the baseline hazard function $h_0(t)$, so that frailty distributions are standardized to $\mathbf{E}(Z) = 1$, the variance parameter $\sigma^2 = \mathbf{V}(Z)$ is interpretable as a measure of heterogeneity across the population in baseline risk. When σ^2 is small, then the values of Z are closely concentrated around one. If σ^2 is large, then values of Z are more dispersed, inducing greater heterogeneity in the individual hazards $Zh_0(t)$ (Wienke, 2006). If there are observed covariates included in the model, equation (3.17) changes to:-

$$h(t \mid X, z) = zh_0(t)e^{\beta' X} \tag{3.18}$$

with $X = (x_1, x_2, \ldots, x_k)'$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ as covariates and regression parameters, respectively. Consequently, a frailty model is a generalization of the well-known proportional hazards model. The proportional hazards model is obtained if the frailty distribution degenerates to Z=1 (Hanagal, 2010).

Let us focus on the main ideas of frailty models. Suppose $S(t \mid Z)$ denotes the survival function of an individual conditional on the frailty Z:

$$S(t \mid Z) = e^{-\int_0^t h_0(s/z)ds} = e^{-Z\int_0^t h_0(s)ds} = e^{-Z\mathrm{H}_0(\mathrm{t})} \tag{3.19}$$

Here $\mathrm{H}_0(\mathrm{t}) = \int_0^t h_0(s)ds$ denotes the cumulative baseline hazard function. The population survival function is obtained from the conditional survival function $S(t \mid Z)$ by integrating out the frailty (Wienke, 2011). It can be viewed as the (unconditional) survival function of an individual randomly drawn from the study population, and corresponds to what can actually be observed:

$$S(t) = \mathrm{E}\{S(t \mid Z)\} = \mathbf{E}\{e^{-Z\mathrm{H}_0(\mathrm{t})}\} = \boldsymbol{\mathcal{L}}\big(\mathrm{H}_0(\mathrm{t})\big) \tag{3.20}$$

where $\boldsymbol{\mathcal{L}}(.)$ denotes the Laplace transform. The derivatives of the Laplace transform can be used to obtain general results about unconditional survival distribution. For example, density and hazard function of the event times and expectation and variance of the frailty can be characterized by the Laplace transform of the frailty distribution and their derivatives (Wienke, 2011). Assume the existence of the following expressions:

$$f(t) = -h_0(t)\boldsymbol{\mathcal{L}}'\big(\mathrm{H}_0(\mathrm{t})\big),$$

$$h(t) = -h_0(t)\frac{\boldsymbol{\mathcal{L}}'(\mathrm{H}_0(\mathrm{t}))}{\boldsymbol{\mathcal{L}}(\mathrm{H}_0(\mathrm{t}))},$$

$$\mathbf{E}\mathrm{Z} = -\boldsymbol{\mathcal{L}}'(0) \quad , \quad \boldsymbol{V}(Z) = \boldsymbol{\mathcal{L}}''(0) - \big(\boldsymbol{\mathcal{L}}'(0)\big)^2 \tag{3.21}$$

where $\boldsymbol{\mathcal{L}}'$ and $\boldsymbol{\mathcal{L}}''$ denote first and second derivatives of the Laplace transform of the frailty distribution . Seeking distributions for the frailty variable Z, it is natural to use those which have explicit Laplace transforms. This will simplify parameter estimation (Hanagal, 2010).

One important problem in the area of frailty models is the choice of the frailty distribution. The frailty distribution most often applied is the gamma distribution (Clayton (1978), Vaupel et al (1979)). In addition to the gamma distribution, distributions such as Gompertz, lognormal, Weibull, and compound Poisson are used to model frailty. Frailty models can be expressed in terms of Laplace transform. Once the Laplace transform of frailty distribution is obtained, it is easy to obtain the estimates of the parameters of frailty models. In this section we discus gamma frailty model only. Details about the other models are available in Wienke (2011).

3.6.1 Univariate Gamma Frailty Model

The gamma distribution has been one of the most widely applied distributions (Greenwood and Yule (1920), Beard (1959), Vaupel et al. (1979), Congdon (1995), dos Santos et al. (1995), Hougaard (2000), Duchateau and Janssen (2008)). From a computational and analytical point of view, it fits very well as a mixture distribution to failure data. It is easy to derive the closed-form expressions of unconditional survival, cumulative density, and hazard functions due to the simplicity of the Laplace transform. The gamma distribution $\Gamma(k,\lambda)$ is a flexible distribution that takes a variety of shapes as k varies: when $k = 1$, it is identical to the well-known exponential distribution; when k is large, it takes a bell-shaped form reminiscent of a normal distribution (Wienke, 2011).

The density of a gamma-distributed random variable with shape k *and* inverse scale parameter λ is given as:

$$g(z) = \frac{\lambda^k z^{k-1} e^{-\lambda z}}{\Gamma(k)} \qquad k > 0, \lambda > 0 \tag{3.22}$$

Consequently, for the Laplace transform it holds that:

$$\mathcal{L}(u) = \frac{1}{\Gamma(k)} \lambda^k \int e^{-uz} z^{k-1} e^{-\lambda z} dz = \left(1 + \frac{u}{\lambda}\right)^{-k} \tag{3.23}$$

The first and second derivatives of the Laplace transform are:

$$\mathcal{L}'(\mathrm{u}) = -\frac{k}{\lambda}\left(1 + \frac{u}{\lambda}\right)^{-k-1}$$

$$\mathcal{L}''(u) = \frac{k(k+1)}{\lambda^2}\left(1 + \frac{u}{\lambda}\right)^{-k-2} \tag{3.24}$$

The mean and variance of Z can be obtained by evaluating the derivatives in (3.21) at = 0 :

$$\mathbf{E}(Z) = \frac{k}{\lambda}, \qquad V(Z) = \frac{k}{\lambda^2}$$

To make sure that the model is identifiable, the restriction $k = \lambda$ is used for the gamma distribution, which results in $\mathbf{E}(Z) = 1$. The variance of the frailty variable is $\sigma^2 = \frac{k}{\lambda^2} = \frac{1}{\lambda}$.

Then the density of a gamma-distributed random variable $Z \sim \Gamma\left(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}\right)$ is given by:

$$g(z) = \frac{1}{\Gamma\left(\frac{1}{\sigma^2}\right)}\left(\frac{1}{\sigma^2}\right)^{\frac{1}{\sigma^2}} z^{\frac{1}{\sigma^2}-1} exp\left(-\frac{z}{\sigma^2}\right) \tag{3.25}$$

The unconditional survival function can be derived by the Laplace transform given in (3.23) as:

$$S(t) = \mathcal{L}\big(\mathrm{H}_0(\mathrm{t})\big) = \big(1 + \sigma^2 \mathrm{H}_0(\mathrm{t})\big)^{-\frac{1}{\sigma^2}} \tag{3.26}$$

This implies the unconditional probability density function:

$$f(t) = h_0(t)\big(1 + \sigma^2 \mathrm{H}_0(\mathrm{t})\big)^{-\left(\frac{1}{\sigma^2}+1\right)}$$

and the unconditional hazard function:

$$h(t) = h_0(t)\big(1 + \sigma^2 \mathrm{H}_0(\mathrm{t})\big)^{-1} \tag{3.27}$$

3.6.2 Estimation of Parametric gamma frailty model

As in the proportional hazards model, a parametric as well as a semiparametric approach is used in estimation. Let us consider the parametric case first. In this approach one can assume a parametric distribution such as Gompertz, Weibull and Exponential to the baseline hazards.

Let the baseline hazard $h_0(t) = \lambda e^{\varphi t}$ be a Gompertz baseline hazard (where $\lambda > 0$ is scale and φ $(-\infty, \infty)$ shape parameters) and frailty follows a gamma distribution $Z \sim \Gamma\left(\frac{1}{\sigma^2}, \frac{1}{\sigma^2}\right)$. Hence the unconditional hazard function (3.27) is given by:

$$h(t) = \lambda e^{\varphi t}\left(\mathbf{1} + \sigma^2 \frac{\lambda}{\varphi}(e^{\varphi t} - 1)\right)^{-\mathbf{1}}$$

For the weibull baseline hazard hazard $h_0(t) = \lambda v t^{v-1}$ (where $\lambda > 0$ and $v > 0$ are the scale and shape parameters, respectively), the unconditional hazard function (3.27) is given by:

$$h(t) = \lambda v t^{v-1}(1 + \sigma^2 \lambda t^{v})^{-\mathbf{1}}$$

For the exponential baseline hazard hazard $h_0(t) = \lambda$ (where λ>0 is scale parameter), the unconditional hazard function (3.27) is given by:

$$h(t) = \lambda(1 + \sigma^2\lambda t)^{-1}$$

As Wienke (2011) described the unconditional likelihood function of the vector of random variables $(T_i, \Delta_i, X_i, Z_i)$ with observation time t_i, covariate vector $X' = (X_{1i}, X_{2i}, \dots, X_{ki})$, censoring indicator δ_i and gamma distributed frailty variable Z_i , is of the form:

$$L(\beta, \theta | Z_1, \dots, Z_n) = \prod_{i=1}^{n} \left(Z_i h_0(t_i; \theta) e^{\beta' X_i}\right)^{\delta_i} e^{-Z_i H_0(t_i;\theta) e^{\beta' X_i}} \quad (3.28)$$

Here θ denotes the vector of parameters of the assumed baseline hazard (e.g. weibull $\theta = (\lambda, v)$). This conditional likelihood function still depends on the unobserved frailty variables. Under the assumption of a gamma-distributed frailty, the random terms can be integrated out using relations (3.27) and (3.26), yielding the likelihood function:

$$L(\beta, \theta, \sigma^2) = \prod_{i=1}^{n} \left(\frac{h_0(t_i;\theta) e^{\beta' X_i}}{1+\sigma^2 H_0(t_i;\theta) e^{\beta' X_i}}\right)^{\delta_i} \left(1 + \sigma^2 H_0(t_i; \theta) e^{\beta' X_i}\right)^{-\frac{1}{\sigma^2}} \quad (3.29)$$

3.6.3 Estimation of Semiparametric gamma frailty model

In the semiparametric frailty model, no assumption about the form of the baseline hazard function is made. This requires new estimation strategies compared to the parametric model. The likelihood function looks similar to the parametric model (3.29), but now the baseline hazard $h_0(t)$ is treated as a nuisance (Wienke, 2011).

- **EM algorithm for univariate gamma frailty**

This algorithm was suggested by Dempster et al. (1977) and is often used in the presence of unobserved data. It was adopted for parameter estimation in frailty models first by Nielsen et. al (1992), Klein (1992), and Guo and Rodriguez (1992).

The EM-algorithm is an iterative parameter estimation procedure used in the analysis of incomplete data to calculate maximum likelihood parameter estimates. Initial ideas of the approach are discussed by Yates (1933). Orchard and Woodbury (1972) suggested the missing information principle and discuss ideas of how to calculate the observed information matrix.

Dempster et al. (1977) introduced the term "EM-algorithm". The issues of convergence were discussed by Wu (1983) and Boyles (1983). Louis (1982) developed techniques for calculation of the observed information matrix.

The EM algorithm iterates between two steps. The E step and M step. In the first, one estimates the expectations of the unobserved frailties based on observed data, and current estimates are obtained. These estimates are then used in the maximization step to obtain new parameter estimates given the estimated frailties. In the gamma frailty model, closed-form expressions exist for the conditional expectations of the frailties in the expectation step.

Furthermore, applying the partial likelihood approach with the estimated frailties as "offset" terms in the M step is easy to perform, which makes the EM algorithm useful. Similar to the expression by Wienke (2011) first we consider the full likelihood with the frailty variables assumed to be observed random variables similar to the event times. We define the full likelihood results from the joint density of $(t_i, \delta_i, Z_i)(i = 1, \dots, n)$ as the product of the conditional and the density of frailties as follows:

$$\begin{aligned} L(\beta, \sigma^2|\mathrm{Z}) &= \prod_{i=1}^{n} f(t_i, \delta_i, Z_i; \beta, \sigma^2) \\ &= \prod_{i=1}^{n} f(t_i, \delta_i, Z_i; \beta) \prod_{i=1}^{n} f(Z_i; \sigma^2) \\ &= L_1(\beta|\mathrm{Z}) L_2(\sigma^2|\mathrm{Z}) \end{aligned} \tag{3.30}$$

$$\text{where } Z = (Z_1, \dots, Z_n) \text{ and } L_1(\beta|\mathrm{Z}) = \prod_{i=1}^{n} \left(Z_i h_0(t_i) e^{\beta' X_i}\right)^{\delta_i} e^{-Z_i H_0(t_i) e^{\beta' X_i}} \tag{3.31}$$

which is the likelihood function of the observed event times conditional on the frailties. The second term $L_2(\sigma^2|\mathrm{Z})$ is given by the probability density of the frailty variables:

$$L_2(\sigma^2|\mathrm{Z}) = \prod_{i=1}^{n} f(Z_i; \sigma^2)$$

If the frailties Z_i were known, the regression parameters β could be estimated by the Cox partial likelihood method rewriting the terms $Z_i e^{\beta' X_i}$ in the form $e^{\beta' X_i + \log(Z_i)}$, using the $\log(Z_i)$ as fixed "offset" values with known regression coefficient of 1. Consequently, the expectation step is needed to get estimates of the frailty values. These estimates are used instead of the unknown frailties in the maximization step to obtain the estimates for the regression parameters (Wienke 2011).

Expectation-Step (E-step): - the unobserved frailty $Z_i (i = 1,..,n)$ of each individual can be estimated by the expression:

$$\mathbf{E}_{(k+1)}(Z_i) = \frac{\frac{1}{\sigma^2{}_{(k)}} + \delta_i}{\frac{1}{\sigma^2{}_{(k)}} + \mathrm{H}_{(k)}(t_i)\mathrm{e}^{\beta'(k)\mathrm{X_j}}} \qquad (3.32)$$

Here $\mathrm{H}_{(k)}(.)$ is a nonparametric estimator of the cumulative baseline hazard based on the current parameter estimates at iteration step k, For example, for the Nelson–Aalen estimator (Nielsen et.al, 1992) we have:

$$\mathrm{H}_{(k)}(t) = \sum_{i:t_i<t} \frac{\delta_i}{\sum_{j\epsilon R(t_i)} \mathbf{E}_{(k)}(Z_j) e^{\beta'(k) X_j}}$$

where $\mathrm{R}(\mathrm{t_i})$ denotes population at risk at time $\mathrm{t_i}$.

Maximization-Step (M-step): - the partial likelihood for the regression parameters in the frailty model is given by:

$$L(\beta|\mathrm{Z}) = \prod_{i=1}^{n} \left(\frac{e^{\beta' X_i + \log(Z_i)}}{\sum_{j\epsilon R(t_i)} Z_j e^{\beta' X_j}} \right)^{\delta_i} \qquad (3.33)$$

The unknown random variables Z_i and $\log(Z_i)$ are now substitutcd by their current expected values (at iteration step k) $\mathbf{E}_{(k)}(Z_i)$ and $\mathbf{E}_{(k)}(\log(Z_i))$:

$$logL(\beta, \sigma^2) = \sum_{i=1}^{n} \delta_i \left[\beta' X_i + \mathbf{E}_{(k)}(\log(Z_i)) - log\left(\sum_{j\epsilon R(t_i)} \mathbf{E}_{(k)}(Z_j) e^{\beta' X_j}\right)\right] \quad (3.34)$$

From this expression, new estimates $\beta_{(k)}$ can be obtained. A new estimate of the frailty parameter $\sigma^2{}_{(k)}$ is derived by maximization of $L_2(\sigma^2|\mathrm{Z})$ also replacing the unknown variables Z_i by their current expected values at iteration step k (Wienke, 2011). Furthermore, as the conditional distribution of Z is gamma, log(Z) has a log-gamma distribution with expectation:

$$\mathbf{E}_{(k+1)}(\log(Z_i)) = \psi\left(\frac{1}{\sigma^2{}_{(k)}} + \delta_i\right) - log\left(\frac{1}{\sigma^2{}_{(k)}} + \mathrm{H}_{(k)}(t_i)\mathrm{e}^{\beta'(k)\mathrm{X_j}}\right)$$

with the derivative of gamma function or digamma function $\psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$

Klein and Moeschberger (2003) provide a modified EM algorithm (KM-EM) for semiparametric PH frailty model and use the information matrix to calculate the variances of the Maximum Likelihood Estimators. For details see Hanagal (2010).

3.7 Bivariate frailty models and Laplace transforms

Let T_1, T_2 be the lifespans of two twins and let Z_1, Z_2 be their individual frailties satisfying:

$$\mathrm{P}(T_i > t_i | Z_1, Z_2) = \mathrm{P}(T_i > t_i | Z_i), i = 1,2 \qquad (3.35)$$

Assume that the twins' lifespans T_1, T_2 are conditionally independent given Z_1, Z_2:

$$\begin{aligned} S(t_1, t_2 | Z_1, Z_2) &= \mathrm{P}(T_1 > t_1, T_2 > t_2 | Z_1, Z_2) \\ &= \mathrm{P}(T_1 > t_1 | Z_1)\mathrm{P}(T_2 > t_2 | Z_2) \\ &= e^{-Z_1 H_{01}(t_1)} e^{-Z_2 H_{02}(t_2)} \end{aligned} \qquad (3.36)$$

where $\mathrm{H_{0j}(t)}$ for $\mathrm{j} = 1,2$ denote the cumulative baseline hazard function. As Iachine (1995) expressed the survival function $S(t_1, t_2)$ can be obtained by averaging the conditional bivariate survival function over the joint distribution of Z_1, Z_2:

$$\begin{aligned} S(t_1, t_2) = \mathbf{E}\big(S(t_1, t_2 | Z_1, Z_2)\big) &= \mathbf{E}\big(e^{-Z_1 H(t_1)} e^{-Z_2 H(t_2)}\big) \\ &= \mathcal{L}(H_{01}(t_1), H_{02}(t_2)) \end{aligned} \qquad (3.37)$$

where $\mathcal{L}(s_1, s_2)$ is the bivariate Laplace transform of Z_1, Z_2 .The model corresponding to the case $Z_1 = Z_2$ is called a shared frailty model (Clayton 1978, Hougaard 1986).

3.7.1 The Concept of Shared Frailty

Shared frailty concept provides multivariate extensions of the traditional univariate frailty model (Vaupel et al. 1979, Lancaster 1979), and it allows mutual dependence of clustered event times to be taken into account in the analysis of event-time data. Survival models for dependent event times are especially useful because they allow more sophisticated examination about the nature of aging, disease, disability, and the mortality processes to be addressed. Such dependence occurs, for example, in event times of related individuals (e.g., family members). The hazard model for each individual in this approach, however, looks exactly the same as in the standard univariate frailty model considered above. The only, but important, difference is that, in a shared frailty model, frailty is defined as a measure of the relative risk that individuals in a group share. Thus the frailty variable is associated with groups of individuals rather than individuals.

The shared frailty approach assumes that all failure times in a cluster are conditionally independent given the frailties. The value of the frailty term is constant over time and common to all individuals in the cluster, and thus it is responsible for creating dependence between event times in a cluster. This dependence is always positive in shared frailty models (Weinke, 2006).

Suppose there are n clusters and that cluster i has n_i observations and associates with the unobserved frailty Z_i $(1 \leq i \leq n)$. The vector X_{ij} $(1 \leq i \leq n, 1 \leq j \leq n_i)$ contains the covariate information of the event time T_{ij} of the j^{th} observation in the i^{th} cluster. The survival times in cluster i $(1 \leq i \leq n)$ are assumed to be independent conditional on the frailty term Z_i, and as Weinke (2006) expressed their hazard functions to be of the form:

$$h(t|X_{ij}, Z_i) = Z_i h_0(t) e^{\beta' X_{ij}} \qquad (3.38)$$

where $h_0(t)$ denotes the baseline hazard function, and β is a vector of fixed effect parameters to be estimated. The frailties Z_i $(i = 1,2, \dots, n)$ are assumed to be independently and identically distributed random variables with density function $g(Z)$. The frailty density depends on unknown parameters to be estimated. Similar to the univariate case, a semiparametric shared frailty model is the one with a nonparametric baseline hazard $h_0(t)$.

Shared frailty models explain correlations within groups (family, litter, or clinic) or for recurrent events facing the same individual. However, this approach does have limitations. First, it forces unobserved factors to be the same within the cluster, which is not generally acceptable. For example, sometimes it may be inappropriate to assume that both partners in a twin pair share all of their unobserved risks. Second, the dependence between survival times within the cluster is based on their marginal distributions of survival times. Third, in most cases, shared frailty will only induce a positive association (Wienke, 2011).

To avoid these limitations, correlated frailty models are being developed for the analysis of multivariate failure time data, in which associated random variables are used to characterize the frailty effect for each cluster. In twin pairs, for example, one random variable is assigned to twin 1 and another to twin 2, so that they are no longer constrained to have a common frailty. These two variables are associated and jointly distributed, therefore, knowing one of them does not automatically imply the other (Wienke, 2011).

3.7.2 The Concept of Bivariate Correlated Frailty

The correlated frailty model is a natural extension of the shared frailty approach on the one hand, and of the univariate frailty model on the other. In the correlated frailty model, the frailties of individuals in a cluster are correlated but not necessarily shared. It enables the inclusion of additional correlation parameters, which then allows the addressing of questions about associations between event times. Furthermore, associations are no longer forced to be the same for all pairs of individuals in a cluster (Wienke, 2011). The conditional survival function in the bivariate case with observed covariates X_j $j(=1,2)$ expressed by Iachine (1995) as:

$$S(t_1, t_2|Z_1, Z_2) = S_1(t_1|Z_1)S_2(t_2|Z_2) = e^{-Z_1 e^{\beta' X_1} H_{01}(t_1)} e^{-Z_2 e^{\beta' X_2} H_{02}(t_2)} \quad (3.39)$$

where Z_1 and Z_2 are two correlated frailties. The distribution of the random vector (Z_1, Z_2) needs to be specified and determines the association structure of the event times in the model.

In the bivariate correlated frailty model, the frailty of each individual in a pair is defined by a measure of relative risk, that is, exactly similar to the univariate case. For two individuals in a pair, frailties are not necessarily the same as they are in the shared frailty model. Here we assume that the frailties are acting multiplicatively on the baseline hazard and that the observations in a pair are conditionally independent, given the frailties. Hence, the hazard of the individual $j(j = 1,2)$ in pair $i(i = 1,2, \dots, n)$ has the form:

$$h(t|x_{ij}, Z_{ij}) = Z_{ij} h_{0j}(t) e^{\beta' X_{ij}} \quad (3.40)$$

where t denotes time, X_{ij} is a vector of observed covariates, β is a vector of regression parameters describing the effect of the covariates X_{ij} , $h_{0j}(\cdot)$ are baseline hazard functions, and Z_{ij} are frailties. Bivariate correlated frailty models are characterized by the joint distribution of a two-dimensional vector of frailties (Z_{i1}, Z_{i2}). If the two frailties are independent, the resulting lifetimes are independent, and no clustering is present in the model or it reduces to univariate frailty. If the two frailties are equal, the shared frailty model is obtained as a special case of the correlated frailty model with correlation one between the frailties (Wienke, 2011).

Similar to univariate and shared frailty models, the choice of the frailty distribution is of great importance for modeling specific features of the joint survival distribution.

3.7.3 Correlated gamma frailty model

This model was introduced by Yashin et al. (1995) and applied to related lifetimes in many different settings, for example, by Pickles et al. (1994), Yashin and Iachine (1995), Yashin et al. (1996), Iachine et al. (1998), Iachine (2002), Petersen (1998). The model has a very convenient representation of the survival function in closed form expressions.

To include heterogeneity in this model, let $Z_i(i = 1,2)$ be the frailties of the two individuals of a twin pair. Assume that their individual hazards are represented by the proportional hazards model as given in (3.40):

$$h(t|\mathrm{X_i}, Z_i) = Z_{ij} h_{0i}(t) e^{\beta' X_i} \qquad (3.41)$$

with a baseline hazard function $h_{0i}(t)$ describing the risk of dying as a function of age.

The following expressions are deduced from the general expression stipulated by Yashin et al (1995) and Wienke (2011) for the application of twins.

Let the lifetimes of the two twin partners be conditionally independent given their frailties Z_1 and Z_2 and, let k_0, k_1, k_2 be some nonnegative real-valued numbers. Then the decomposition of Z_1 *and* Z_2 becomes:

$$Z_1 = Y_0 + Y_1$$
$$Z_2 = Y_0 + Y_2 ,$$

where Y_0, Y_1 *and* Y_2 are independent gamma-distributed same scale λ random variables with density:

$$g(Y_j) = \frac{\lambda^{k_j} z^{k_j - 1} e^{-\lambda Y_j}}{\Gamma(k_j)} \qquad k_j > 0, \lambda > 0, \; j = 0,1,2 \qquad (3.42)$$

Obviously, Z_1 and Z_2 are correlated in view of the shared part of frailty Y_0 in both Z_1 and Z_2 . To force Z_1 and Z_2 to have the same distribution we assume that shape parameters k_1*and* k_2 for the distributions of Y_1 *and* Y_2 are the same $(k_1 = k_2)$. This condition is relevant for twin studies when there is no reason to argue different distributions of frailty for two twins, and can be omitted for other applications. It can be shown that under these assumptions frailties Z_1 and Z_2 are gamma-distributed correlated random variables (Yashin et al, 1995);

$$Z_1 \sim \Gamma(k_0 + k_1, \lambda) \qquad (3.43)$$
$$Z_2 \sim \Gamma(k_0 + k_1, \lambda)$$

Furthermore, one can employ the standard assumption that the mean frailty of individuals is one (at the beginning of the follow-up), which means that:

$$\mathbf{E}(Z_1) = \mathbf{E}(Z_2) = \frac{k_0+k_1}{\lambda} = 1$$

and the common variance :

$$\mathbf{V}(Z) = \mathbf{V}(Z_1) = \mathbf{V}(Z_2) = \frac{1}{\lambda} = \sigma_z^2$$

This leads to the correlation coefficient of Z_1 and Z_2 given by:

$$\rho_z = \frac{\mathbf{cov}(Z_1,Z_2)}{\sqrt{\mathbf{V}(Z_1)\mathbf{V}(Z_2)}} = \frac{k_0}{k_0+k_1} \tag{3.44}$$

$$\text{Let } \lambda = \frac{1}{\sigma_z^2}, k_0 = \frac{\rho_z}{\sigma_z^2}, k_1 = k_2 = \frac{1-\rho_z}{\sigma_z^2} \tag{3.45}$$

then

$$Z_j \sim \Gamma\left(\frac{1}{\sigma_z^2}, \frac{1}{\sigma_z^2}\right) \; for\; j = 1,2$$

Now we can derive the unconditional survival function applying the Laplace transform of gamma-distributed random variables. Hence,

$$\begin{aligned} S(t_1,t_2) &= \mathbf{E}S(t_1,t_2|Z_1,Z_2) = \mathbf{E}S_1(t_1|Z_1)S_2(t_2|Z_2) = \mathbf{E}e^{-Z_1H_{01}(t_1)}e^{-Z_2H_{02}(t_2)} \\ &= \mathrm{E}e^{-(Y_0+Y_1)H_{01}(t_1)}\mathrm{e}^{-(Y_0+Y_2)H_{02}(t_2)} \\ &= \left(1+\sigma_z^2\mathrm{H}_{01}(\mathrm{t}_1)+\sigma_z^2\mathrm{H}_{02}(\mathrm{t}_2)\right)^{-\frac{\rho_z}{\sigma_z^2}} \\ &\quad \times \left(1+\sigma_z^2\mathrm{H}_{01}(\mathrm{t}_1)\right)^{\frac{-1+\rho_z}{\sigma_z^2}}\left(1+\sigma_z^2\mathrm{H}_{02}(\mathrm{t}_2)\right)^{\frac{-1+\rho_z}{\sigma_z^2}} \end{aligned} \tag{3.46}$$

Using the relationship (3.26), the above representation can be transformed to a semiparametric form of the correlated gamma frailty model as:

$$s(t_1,t_2) = \frac{S_1(t_1)^{1-\rho_z}S_2(t_2)^{1-\rho_z}}{\left(S_1(t_1)^{-\sigma_z^2}+S_2(t_2)^{-\sigma_z^2}-1\right)^{\frac{\rho_z}{\sigma_z^2}}} \tag{3.47}$$

where $S(t)$ denotes the marginal univariate survival function, which is assumed to be equal for both partners in a twin pair. It is estimated by using either univariate parametric or non parametric methods.

The bivariate survival function without frailty and the shared gamma frailty model are special cases of the correlated gamma frailty model (Hanagal, 2010). The special cases and the corresponding models are expressed as:

$$S(t_1,t_2)=\begin{cases}S_1(t_1)S_2(t_2), & if\ \sigma_z^2=0\ or\rho_z=0\\ \left(S_1(t_1)^{-\sigma_z^2}+S_2(t_2)^{-\sigma_z^2}-1\right)^{-\frac{1}{\sigma_z^2}}, & if\ \rho_z=1\ and\ \ \sigma_z^2>0\end{cases} \tag{3.48}$$

The first model $S_1(t_1)S_2(t_2)$ is a bivariate survival function without frailty. The absence of frailty implies the paired observations are independent. The second model is the shared gamma frailty model.

The correlated gamma frailty model with observed covariates is a simple extension of (3.47):

$$S(t_1,t_2|\mathrm{X}_1,\mathrm{X}_2)=\frac{S_1(t_1|\mathrm{X}_1)^{1-\rho_z}S_2(t_2|\mathrm{X}_2)^{1-\rho_z}}{\left(S_1(t_1|\mathrm{X}_1)^{-\sigma_z^2}+S_2(t_2|\mathrm{X}_1)^{-\sigma_z^2}-1\right)^{\frac{\rho_z}{\sigma_z^2}}} \tag{3.49}$$

where $\mathrm{X}_1$and X_2 are the covariate vectors of the partners in the pair. The marginal survival functions are assumed to be identical in applications to twin data such that:

$$S(t|X)=S_1(t_1|\mathrm{X}_1)=S_2(t_2|\mathrm{X}_2)$$

$$S(t|X)=\left(1+\sigma_z^2\mathrm{H}_0(\mathrm{t})e^{\beta\cdot X}\right)^{-\frac{1}{\sigma_z^2}} \tag{3.50}$$

3.7.4 Parameter estimation in Bivariate Correlated Gamma Frailty Model

Two major methods can be used for parameter estimation. The first one is "semiparametric" estimation. It does not assume a parametric specification of the underlying hazard. It works with observed covariates that can be done by the EM-algorithm for the analysis of the correlated gamma-frailty model developed by Iachine (1995) or Peterson et al (1996). The second one is the "parametric" estimation. It uses a parametric specification of the underlying hazard and bivariate frailty distribution in the model for the maximum likelihood estimation algorithm.

- **Semiparametric Estimation**

A version of the EM-algorithm for the analysis of the correlated gamma-frailty model was developed by Iachine (1995) and Petersen et al (1996). Parner (1998) showed the consistency and asymptotic normality of the nonparametric maximum likelihood estimators of the correlated

gamma-frailty model parameters with observed covariates. Similar to univariate EM estimation, in each of the iterations the parameter estimates obtained on the previous iteration are updated. Under some standard regularity conditions, the sequence of estimates converges to the maximum likelihood estimates of the parameters (Iachine, 1995).

Let $(t_{ij}, \delta_{ij}, X_{ij})$ be the vector of observed survival times t_{ij}, censoring information ($\delta_{ij} = 0,1$; $\delta_{ij} = 0$ indicates right censoring) and observed values of covariate vector X_{ij} for the j^{th} individual in the i^{th} pair ($j = 1,2; i = 1,2, \dots, n$). Assume identical baseline hazard for the two individuals in a twin pair ($h_{01} = h_{02} = h_0$). Iachine (1995) illustrate that if the frailty components y_{ij} in (3.42) are observed along with $(t_{ij}, \delta_{ij}, X_{ij})$, then the correlated gamma frailty model with observed covariates has a proportional hazard structure. The structure allows us to combine Cox's regression and maximum likelihood techniques to obtain parameter estimates of $\sigma_z^2, \rho_z, \beta$ and to calculate a semiparametric estimate of $H(t)$. The details about the whole procedure are available in Iachine (1995) .

- **Parametric Estimation**

In "parametric" estimation approach equation (3.46) can be used to form the likelihood function of the data. The likelihood function under independent and non informative right censoring without truncation can be expressed in terms of the bivariate survival function by:

$$\prod_{i=1}^{n} S_{t_{i1},t_{i2}}(t_{i1}, t_{i2})^{\delta_{i1}\delta_{i2}} S_{t_{i1}}(t_{i1}, t_{i2})^{\delta_{i1}(1-\delta_{i2})} S_{t_{i2}}(t_{i1}, t_{i2})^{(1-\delta_{i1})\delta_{i2}} S(t_{i1}, t_{i2})^{(1-\delta_{i1})(1-\delta_{i2})} \quad (3.51)$$

The detail about how the above likelihood function is derived is available in weink (2010).

The details of maximum likelihood parameter estimation procedure for weibull baseline hazard are shown in Annex B. One can follow a similar procedure for gompertz and exponential baseline hazard functions.

To obtain the standard error of the parameters, the observed information matrix is computed, and an approximate Fisher information matrix is obtained by inserting the parameters estimates into it. The observed information matrix for a vector of parameters say $\Omega = (\sigma_z^2, \rho_z, \beta, \theta)$ is given by:

$$J(\theta) = -\frac{\partial^2}{\partial \Omega_i' \partial \Omega_j} L(\Omega)$$

Here θ denotes the vector of parameters of the assumed baseline hazard (e.g. weibull $\theta = (\lambda, v)$). The asymptotic normality of the maximum likelihood estimators in bivariate correlated gamma frailty model justifies the calculation of asymptotic confidence intervals and the application of tests like the likelihood ratio test (see Giard, 2001).

3.8 Tests of hypothesis in frailty model

a) Test of heterogeneity

The hypothesis to be tested are $H_o: \sigma_z^2 = 0 \; versus \; H_1: \sigma_z^2 > 0$, where σ_z^2 is the variance of the frailty (unobserved heterogeneity). The main problem here is that the heterogeneity parameter is on the boundary of its parameter space under H_o and standard methods cannot be applied. Approximation of the likelihood ratio test statistic using x^2- distribution with one degree of freedom is too conservative (Self and Liang, 1987). As a result, instead of using the usual x^2 – distribution, a mixture of a $0.5x_0^2 \; and \; 0.5x_1^2$ distribution should be used (Claeskens et al. 2008). Note that in semiparametric method of parameter estimation, if $\sigma_z^2 = 0$ the frailty model is reduced to the usual Cox model. Thus, this test can also be used to test the appropriateness of the Cox model.

b) Test of shared frailty

The shared frailty model is the special case of correlated frailty models when $\rho_z = 1$. Thus the hypothesis to be tested are $H_o: \rho_z = 1 \; versus \; H_1: \rho_z < 1$, where ρ_z is the correlation coefficient of the frailties Z_1 and Z_2 . Similar to the test of heterogeneity, the main problem here is that the correlation parameter is on the boundary of its parameter space under H_o and standard methods cannot be applied. As a result, instead of using the usual x^2 –distribution, a mixture of a $0.5x_0^2 \; and \; 0.5x_1^2$ distribution should again be used (Claeskens et al. 2008).

Chapter four

4. Results and discussion

4.1 Descriptive statistics

A total of 1178 or (589 pairs of) twin child deliveries were recorded in the 2011 Ethiopia Demographic and Health survey. The data on 1056 children or (528 pairs of twin) born before January 01, 2008 were analyzed in this study. The overall information on censoring and covariates that are included in this study are presented below.

Table 4. 1: Summary on censoring status of the study population

Status	Number of pairs
both twins dead	143
one twin alive, cotwin dead	145
both twins alive	240
all pairs together	528

Table 4. 2: Summary results of the covariates included in this study

Covariates	Category	Number of children (%)	Number of death
Residence	rural	852 80.7%	371
	urban	204 19.3%	60
Sex	female	520 49.3%	195
	male	536 50.7%	236
Age of mother at birth	below18 years	174 16.5%	88
	between 18-35 years	816 77.3%	323
	above 35 years	66 6.2%	20
Preceding birth interval	below 18 months	154 14.6%	81
	between 18-24 months	132 12.5%	58
	above 24 months or 1st born	770 72.9%	292
Birth order	first born	170 16.1%	86
	between (2 -4)	584 55.3%	230
	5 and above	302 28.6%	115
Previous child status	dead	208 19.7%	105
	alive or first child	848 80.3%	326

The summary results in Table 4.2 shows that 80.7% of the study population lived in rural areas, 19.3% resided in urban areas. When we categorize the study population using the variable sex 50.7% of the study population are males and the remaining 49.3% are females. Among the total 536 males in the study there are 236 deaths, whereas among the 520 females in the study 195 female children are dead. The variable age of mother at birth is categorized in to three groups. In the first category, below 18 years, there are 16.5% of the study population, while 6.2% of the children were born from mothers whose age exceeds 35 years. The remaining 77.3% children belong to mothers whose age at birth was between 18 - 35 years. A total of 886 children (83.9%) of the study population have younger sibling whereas the remaining 16.1% are first born children. About 14.6% of the children were born before their younger sibling reaches the age 18 months and 12.5% of the children were born when their younger sibling's age was between 18 - 24 months. The remaining 72.9% of the study population have no younger sibling or their younger sibling's age exceed 24 months. The younger sibling for 19.7% of the study population have died.

The graph of the estimate of overall Kaplan-Meier survivor function is given in Figure 4.1. As one can see that there is high mortality of twins before they reach 52 weeks (one year). To include covariates in survival regression models, particularly for categorical variables, we need look at the Kaplan-Meier estimates of the group-specific survivorship functions. Hence separate graphs of the estimates of the Kaplan-Meier survivor functions for all covariates were plotted in order to see whether there were differences in survival experiences between different categories of covariates (see Annex A). The graphs indicate that there are differences between the various categories.

To test if there are statistically significant differences among the survival experience of the different groups of the covariates, the log rank test was performed. As we can see in Table 4.3 below, the results of the log rank test indicate that there are significant differences between the survival experience of different categories for all the covariates.

Figure 4. 1:The Kaplan -Meier estimate of $S(t)$

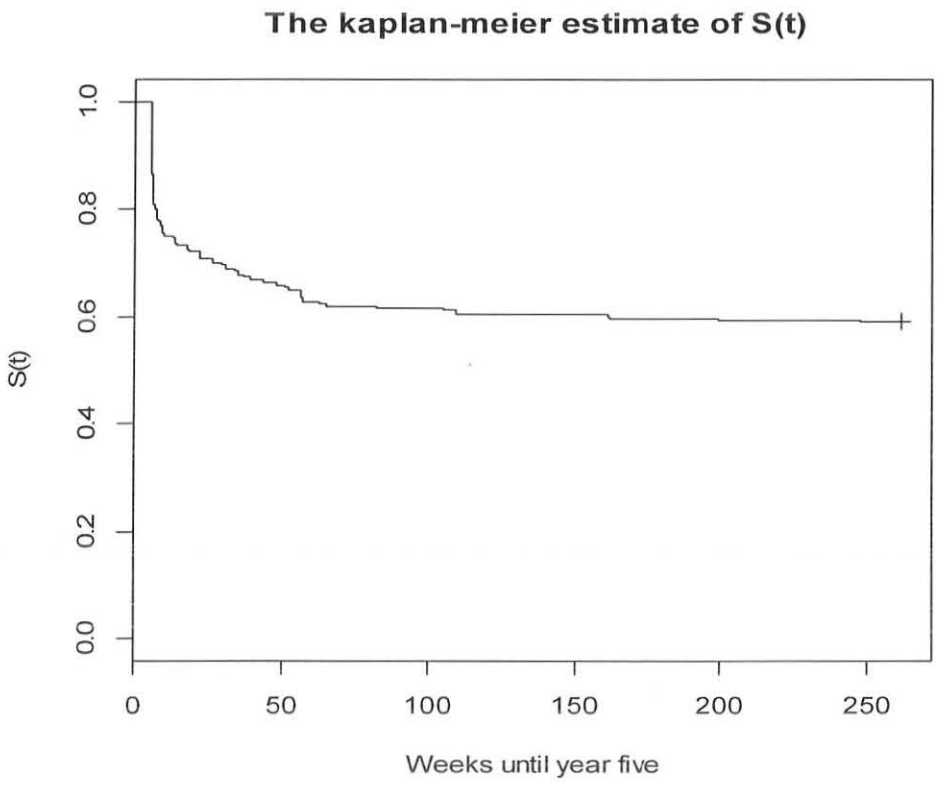


Table 4. 3: Log-Rank Test for covariates

Variables	Value of the test statistic	P value
Residence	14.1	0.000177
Sex	4.4	0.0358
Age of mother at birth	10	0.00687
Preceding birth interval	12.2	0.00229
Birth order	10.2	0.00616
Previous child status	9.6	0.00199

4.2 Results of the Cox proportional hazards model

Before proceeding to more complicated models, we first fit a univariate Cox proportional hazards regression model for every potential risk factor. The likelihood ratio test is considered for each univariate Cox PH model.

Table 4. 4 :Univariate Cox proportional hazards model result

Covariates	DF	$\hat{\beta}$	SE($\hat{\beta}$)	Wald x^2	P value	-2LOGL
Residence	1	0.511	0.139	15.2	<.0001*	5790.655
Sex	1	-0.201	0.0968	4.31	0.0378*	5801.513
Age of mother at birth	1	-0.314	0.101	9.49	0.00207*	5796.34
Preceding birth interval	1	-0.206	0.0612	10.7	0.00109*	5795.152
Birth order	1	-0.173	0.0747	5.35	0.0207*	5800.472
Previous child status	1	-0.342	0.112	8.75	0.0031*	5797.079

* variable is significant at the 5% level of significance

As can be seen from the above Table 4.4 the result of the univariate Cox proportional hazards model shows that all covariates are significant at 5 % level.

Then we fit the multivariate Cox PH model including all the potential risk factors. The result is given in Table 4.5 below. The result shows that all covariates are significant at 5 % level. Furthermore, the results of the multivariate Cox PH model for the categories of the variables in Table 4.5 below show that the categories of the variable age of mother at birth are found to be insignificant at 5 % level.

The shared gamma frailty model is fitted to assess the random effect shared by the twins. In semiparametric estimation, introducing a gamma frailty term in the model is similar to adding one more covariate with a known regression coefficient of one. In other words, if the frailty variable is known, it is just a Cox PH model with an additional covariate with a known regression coefficient. The result is given in Table 4.6.

Table 4. 5: Results of the Cox proportional hazards model

Covariates	$\hat{\beta}$	SE($\hat{\beta}$)	z	P value	**HR**	95% CI for **HR** LCL	UCL
Residence	0.608	0.1424	4.27	<.0001*			
(rural)	0.6409	0.1433	4.473	<.0001*	1.898	1.433	2.513
Ref (urban)					1.00		
Sex	-0.209	0.0975	-2.14	0.032*			
(female)	-0.1987	0.0977	-2.035	0.042*	0.819	0.677	0.992
Ref (male)					1.00		
Age of mother at birth	-0.258	0.1146	-2.25	0.025*			
(below18 years)	0.4875	0.2713	1.797	0.072	1.628	0.956	2.770
(between 18-35 years)	0.2568	0.2405	1.068	0.29	1.292	0.806	2.071
Ref (above 35 years)					1.00		
Preceding birth interval	-0.213	0.0660	-3.23	0.0012*			
(below 18 months)	0.4439	0.1383	3.208	0.0013*	1.558	1.188	2.044
(between 18-24 months)	0.2799	0.1491	1.878	0.060	1.323	0.987	1.772
Ref (> 24 month or 1stborn)					1.00		
Birth order	-0.240	0.0903	-2.66	0.0079*			
(first born)	0.5875	0.1746	3.365	0.0007*	1.799	1.278	2.533
(between 2 -4)	-0.0185	0.1217	-0.152	0.88	0.981	0.773	1.246
Ref (5 and above)					1.00		
Previous child status	-0.236	0.1178	-2.01	0.045*			
(dead)	0.3035	0.1217	2.494	0.013*	1.346	1.06	1.719
Ref (alive or first child)					1.00		
Log likelihood(model)	-2871.021						
Log likelihood(null)	-2902.913						

* variable is significant at the 5% level of significance

The result of the shared gamma frailty model in Table 4.6 below shows that the variables residence and preceding birth interval found to be significant at 5% level. An estimate of the

heterogeneity parameter σ_z^2 is 1.62. The likelihood ratio test statistic of H_0: $\sigma_z^2 = 0$ *versus* H_1: $\sigma_z^2 > 0$ is found to be significant at the 1% level.

Table 4. 6 :Results of the Shared gamma frailty model

Covariates	$\hat{\beta}$	SE($\hat{\beta}$)	Chi square	P value	**HR**	95% CI for **HR** LCL	UCL
Residence	0.801	0.211	14.42	0.00015*			
(rural)	0.863	0.212	16.50	<.0001*	2.370	1.563	3.59
Ref (urban)					1.00		
Sex	-0.143	0.129	1.22	0.27			
(female)	-0.136	0.128	1.13	0.29	0.873	0.678	1.12
Ref (male)					1.00		
Age of mother at birth	-0.364	0.199	3.37	0.0667			
(below18 years)	0.823	0.433	3.61	0.058	2.277	0.974	5.32
(between 18-35 years)	0.642	0.365	3.10	0.078	1.901	0.929	3.89
Ref (above 35 years)					1.00		
Preceding birth interval	-0.279	0.113	6.12	0.013*			
(below 18 months)	0.633	0.234	7.32	0.0068*	1.883	1.191	2.98
(between 18-24 months)	0.390	0.243	2.56	0.110	1.476	0.916	2.38
Ref (> 24 month or 1st born)					1.00		
Birth order	-0.241	0.143	2.83	0.092			
(first born)	0.724	0.296	5.97	0.015*	2.063	1.154	3.69
(between 2 -4)	-0.252	0.200	1.59	0.210	0.777	0.525	1.15
Ref (5 and above)					1.00		
Previous child status	-0.376	0.201	3.50	0.061			
(dead)	0.483	0.203	5.68	0.017*	1.621	1.090	2.41
Ref (alive or first child)					1.00		
σ_z^2	1.62	0.1961		<.0001*			
Log likelihood(model)	-2500.103						
Log likelihood(null)	-2902.913						

* variable is significant at the 5% level of significance

4.2.1 Results of model diagnostics

After a Cox PH model with or without frailty is fitted, the adequacy of the model, including the PH assumption and the goodness of fit, needs to be assessed.

- **Assessing proportionality assumption**

In order to check the proportionality assumption, we fit Cox PH with all covariates and time-dependent interaction terms for all covariates simultaneously. The result in Table 4.7 shows that the coefficients of all time-dependent interaction terms were insignificant indicating that the proportional hazards assumption is not violated for all covariates.

Table 4. 7: The result of the fitted Cox PH model with covariates and covariates*time

covariates	coefficient	Standard Error	Z calculated	P value
Residence	4.47e-02	1.76e-01	0.2543	0.80
Sex	-2.12e-02	1.16e-01	-0.1827	0.86
Age of mother at birth	-4.73e-03	1.44e-02	-0.3293	0.74
Preceding birth interval	-1.50e-03	3.20e-03	-0.4679	0.64
Birth order	1.10e-02	3.14e-02	0.3514	0.73
Previous child status	8.14e-03	1.39e-01	0.0586	0.95
Residence*Time	-3.83e-03	3.49e-03	-1.0949	0.27
Sex*Time	1.00e-03	3.11e-03	0.3231	0.75
Age of mother at birth*Time	9.70e-05	2.92e-04	0.3326	0.74
Preceding birth interval *Time	4.89e-05	7.74e-05	0.6318	0.53
Birth order*Time	1.29e-05	7.79e-04	0.0165	0.99
Previous child status *Time	-1.75e-03	3.68e-03	-0.4746	0.64

- **Assessment of goodness of fit**

To assess the goodness of fit of the two fitted models, for each model we plot the Cox-Snell residuals ($r_{ci} = \exp(\hat{\beta}' X_i) * \widehat{H_0}(t_i)$) against the cumulative hazard of the Cox-Snell residuals. The cumulative hazard of the Cox-Snell residual is obtained by using the Kaplan-Meier

estimates of the survivor function. If the observed survival time t_i is censored, the corresponding r_{ci} is also censored. If the fitted model is correct, the plot gives a straight line with unit slope and zero intercept.

As can be seen from Figure 4.2 and Figure 4.3 below, there are deviations from the straight line in the right-hand tails of the two plots. However, deviations from the straight line in the right-hand tail of the distribution could be due in part to uncertainty about the estimates $\hat{\beta}$ *and* $\widehat{\mathrm{H}_0}(t_i)$, since in this area the baseline hazard is more variable due to the reduced effective sample size caused by prior failures and censoring (Mara and Jong Sung, 2004). Thus, we can say that both models fit the data well.

Figure 4. 2: Cumulative hazard plot of the Cox-Snell residual for Cox PH model

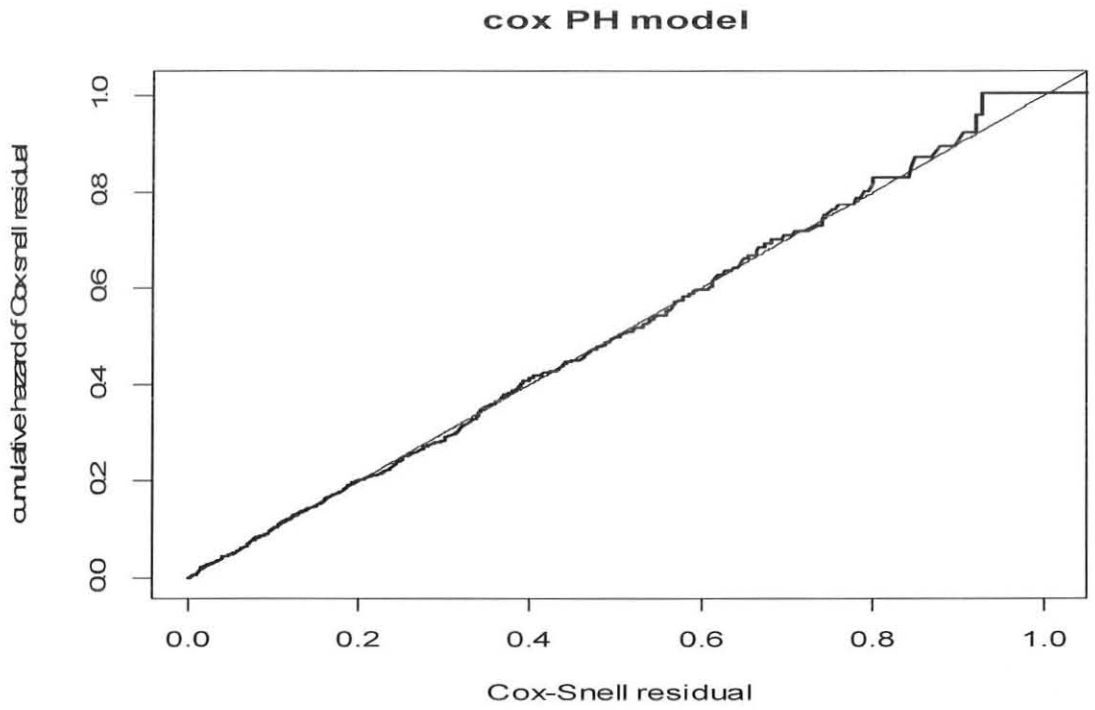
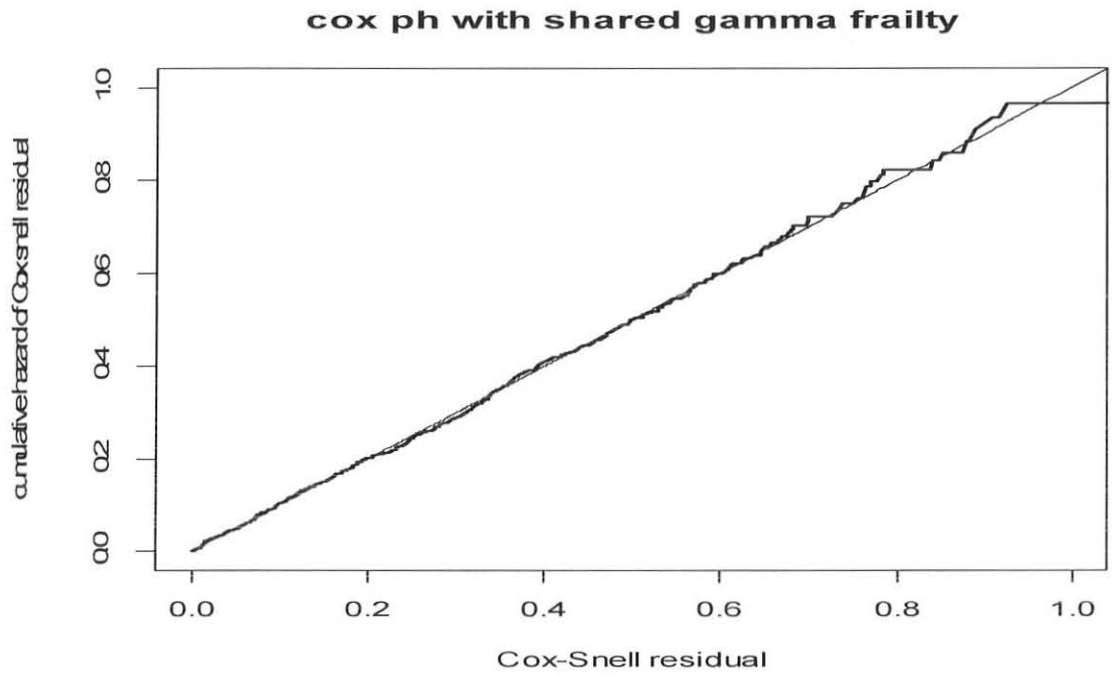


Figure 4. 3: Cumulative hazard plot of the Cox-Snell residual for shared gamma frailty model



4.2.2 Comparison of Cox PH versus shared gamma frailty model

From Table 4.8 below we can see that the shared gamma frailty model has a larger log likelihood and minimum AIC and BIC values than the Cox PH model, indicating that this model fits the data better than the Cox PH model which did not take into account the shared random effect. Since Cox PH model is nested within shared gamma frailty model we conduct the likelihood ratio test for the model. Furthermore, we have already proved that the heterogeneity parameter σ_z^2 is significant. Thus, the shared gamma frailty model is a better fit.

Table 4. 8: Comparison of Cox PH and Shared frailty model

Model	**Log-likelihood (null)**	**Log-likelihood (model)**	DF	AIC	BIC
Cox PH	-2902.913	-2871.021	9	5760.041	5804.702
Shared frailty	-2902.913	-2500.103	9	5018.207	5062.867

4.2.3 Interpretations of the results

The results of the analysis using shared gamma frailty model in Table 4.6 shows that the variables place of residence and preceding birth interval are significantly associated with twin under-five mortality. Furthermore, the categories of the covariates birth order and previous child status are found significant. Here we interpret the results as follows.

The estimated hazard ratio of a twin child (i.e, a child of twin birth) born in rural areas is 2.370 (95% CI: 1.563- 3.59) implying that the risk of dying for a child of twin birth in rural areas is 137% higher than that in urban areas controlling for the other covariates in the model. This figure can be as low as 56.3% and as high as 259% with 95 % confidence.

The estimated hazard ratio of a child of twin birth who was born before the younger sibling reaches the age of 18 months is 1.883 (95% CI: 1.191- 2.98). Thus, the hazard rate of a child of twin birth who was born before the younger sibling reaches the age of 18 months is 1.88 times than that who either has no younger sibling or born when the younger sibling's age exceeds 24 months (reference group) controlling for other covariates in the model. The confidence interval indicates that the risk of dying can be as low as 1.191 times and as high as 2.98 times than that of the reference group.

The estimated hazard ratio of a first born child of twin birth is 2.063 (95% CI: 1.154-3.69). This shows that a child of first born twins has a 106.3% higher risk of dying than a twin child whose birth order is five and above (reference group) controlling for other covariates. This figure can be as low as 15.4% and as high as 269% with 95 % confidence.

Children of twin birth whose previous sibling died experience a lower survival chance compared to those whose previous sibling is alive or first born (reference group). The estimated hazard ratio is 1.621 (95% CI: 1.090, 2.41). This shows that the risk of dying for a child of twin birth whose younger sibling had died is 62.1% higher than the reference group.

4.3 Results of the parametric proportional hazards model

Fitting a parametric proportional hazards model requires assumptions about the parametric form of the underlying cumulative hazard function. To this end, we critically observe the plots of semi-parametrically estimated cumulative baseline hazards of the above two models. As one can see from the plots Figure 4.4 and Figure 4.5, both estimated cumulative baseline hazards resemble the cumulative baseline hazard of Gompertz with a negative shape parameter and Weibull with shape parameter taking a value between (0,1). As a result we fit a parametric PH (Gompertz and Weibull) model for this data set.

Figure 4. 4: Plot of time versus cumulative baseline hazard for Cox PH model

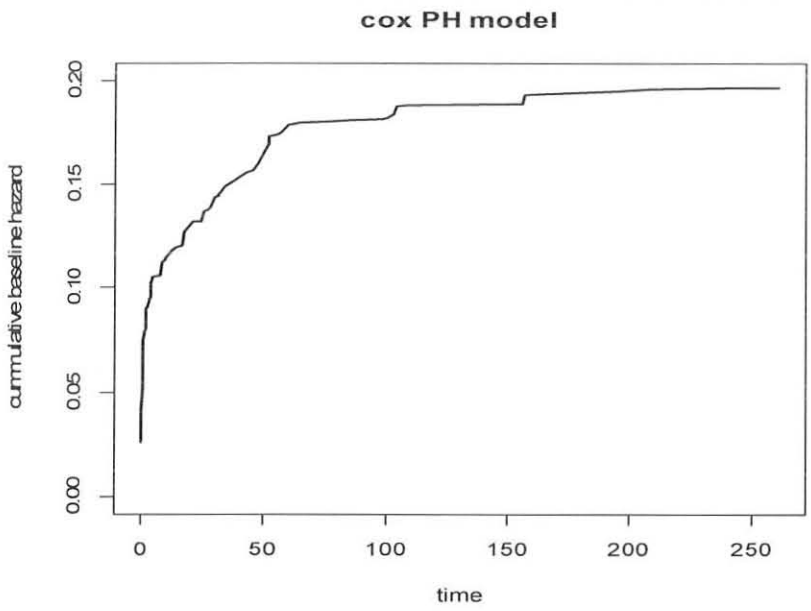
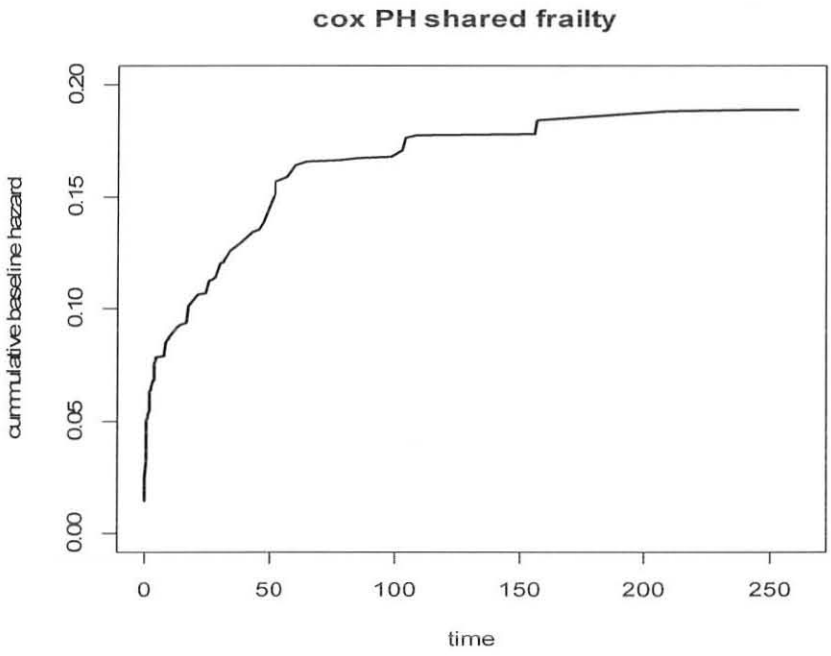


Figure 4. 5: Plot of time versus cumulative baseline hazard for shared gamma frailty model



4.3.1 Results of the parametric bivariate proportional hazard models without frailty

Firstly we fit:

$$S(t_1, t_2|X_1, X_2) = S(t_1|X_1)S(t_2|X_2)$$

which means a bivariate PH model without frailty ($\sigma_z^2 = \rho_z = 0$) assuming the paired observations are independent. Provided that the parametric baseline hazard is reasonable or fits well, the estimated regression coefficients of this model are expected to be similar to those coefficients obtained by Cox PH model. This is justified by a simulation study given in Annex C.

The results in Table 4.9 below indicate that the Gompertz model seems a better fit to the data (based on the log likelihood, AIC and BIC) compared to the weibull model and should be preferred.

Table 4. 9: The result of the fitted bivariate parametric PH without frailty model

covariates	Without frailty Weibul baseline hazard			Without frailty Gompertz baseline hazard		
	Coef	SE	p value	Coef	SE	p value
Residence						
rural	0.68199	0.14241	3.432e-07*	0.67432	0.13657	4.512e-07*
Ref (urban)						
Sex						
female	-0.21409	0.09771	0.02905*	-0.20869	0.09730	0.03231*
Ref (male)						
Age of mother at birth						
below18 years	0.50984	0.26665	0.05681	0.49217	0.24704	0.06025
between 18-35 years	0.23408	0.23647	0.31657	0.25867	0.21356	0.26724
Ref (above 35 years)						
Preceding birth interval						
below 18 months	0.48158	0.13885	0.00078*	0.44401	0.13867	0.00189*
between 18-24 months	0.30329	0.14935	0.05849	0.29254	0.14899	0.05665
Ref (> 24 months or 1st born)						
Birth order						
first born	0.62956	0.17601	0.00038*	0.61080	0.17534	0.00057*
between 2 -4	-0.00840	0.12309	0.94561	-0.03434	0.12289	0.78086
Ref (5 and above)						
Previous child status						
dead	0.32928	0.12189	0.00811*	0.31014	0.12196	0.01265*
Ref (alive or first child)						
Scale (λ)	0.021709	0.00593		0.00702	0.00147	
Shape(φ)	0.402106	0.01747		-0.03536	0.00181	
Log likelihood	-2652.262			-2475.139		
AIC	5326.525			4972.277		
BIC	5381.109			5026.862		

*the variable is significant at 5% level of significance

4.3.2 Results of the parametric bivariate proportional hazards models with gamma frailty

To assess the random effect shared by two individuals in a pair, we fit parametric shared frailty models. The result in Table 4.10 also shows that the Gompertz model is a better fit to the data compared to the weibull model and should be preferred.

Table 4. 10: The results of the fitted parametric PH shared frailty models

covariate	Shared gamma frailty with Weibul baseline hazard			Shared gamma frailty with Gompertz baseline hazard		
	Coef	SE	p- value	Coef	SE	p value
Residence						
rural	1.02762	0.23775	7.69e-05*	0.93114	0.20423	3.42e-05*
Ref (urban)						
Sex						
female	-0.15578	0.14042	0.26968	-0.14650	0.13063	0.26530
Ref (male)						
Age of mother at birth						
below18 years	1.00432	0.46967	0.05500	0.84095	0.39203	0.06331
between 18-35 years	0.81635	0.37577	0.06533	0.68324	0.30942	0.07525
Ref (above 35 years)						
Preceding birth interval						
below 18 months	0.63578	0.27563	0.01993*	0.56138	0.24162	0.01980*
between 18-24 months	0.48228	0.28915	0.09223	0.40711	0.25297	0.10649
Ref (> 24 months or 1stborn)						
Birth order						
first born	0.77494	0.35662	0.03028*	0.70549	0.30991	0.02388*
between 2 -4	-0.31278	0.23905	0.18859	-0.29344	0.21001	0.16106
Ref (5 and above)						
Previous child status						
dead	0.55315	0.24096	0.02078*	0.47913	0.21075	0.02261*
Ref (alive or first child)						
Scale (λ)	0.00738	0.00265		0.00519	0.00133	
Shape(φ)	0.62707	0.03045		-0.02594	0.00177	
σ_z^2	2.67447	0.31862	1.469e-45*	1.85274	0.24259	1.193e-31*
Log likelihood	-2552.603			-2407.317		
AIC	5129.205			4838.634		
BIC	5188.752			4898.181		

*indicate the variable is significant at 5% level of significance

An estimate of the heterogeneity parameter σ_z^2 is 1.85274 (SE=0.24259). The likelihood ratio test statistics for testing Ho: $\sigma_z^2 = 0$ *versus* H_1: $\sigma_z^2 > 0$ is significant at 1% level. Thus, we cannot ignore the correlation between the pair of twins. In other words, the shared gamma frailty model with gompertz baseline hazard is more appropriate than that without frailty.

Finally we fit parametric correlated gamma frailty models using the two specified cumulative baseline hazards. The results indicate that the correlated gamma frailty model with Gompertz baseline hazard has a larger log likelihood value (-2406.571) compared to weibull (-2552.118). The result of the fitted parametric correlated gamma frailty with Gompertz cumulative baseline hazard is given in Table 4.11.

Table 4. 11: The result of the fitted parametric (gompertz baseline) correlated gamma frailty model

covariates	**Correlated gamma frailty with Gompertz baseline hazard**		
	Coef	**SE**	**p value**
Residence			
rural	0.97159	0.21795	4.768e-05*
Ref (urban)			
Sex			
female	-0.16738	0.13947	0.23182
Ref (male)			
Age of mother at birth			
below18 years	0.83307	0.41641	0.08673
between 18-35 years	0.68201	0.32867	0.09876
Ref (above 35 years)			
Preceding birth interval			
below 18 months	0.55123	0.25572	0.03200*
between 18-24 months	0.38309	0.26742	0.15302
Ref (> 24 months or 1stborn)			
Birth order			
first born			
between 2-4	0.73188	0.32713	0.02656
Ref (5 and above)	-0.27521	0.22109	0.21391
Previous child status			
dead	0.48915	0.22379	0.02854*
Ref (alive or first child)			
Scale (λ)	0.00525	0.00143	
Shape(φ)	-0.02397	0.00237	
σ_z^2	2.26442	0.43719	5.629e-32*
ρ_z	0.87251	0.09369	0.11088
Log likelihood	-2406.571		
AIC	4839.141		
BIC	4903.65		

*the variable is significant at 5% level of significance

Usually correlated gamma frailty model is fitted to assess genetic effect. To this end, we need the zygosity information that would enable us to compare the correlation between MZ and DZ twins. Moreover, this model is also used to test hypotheses about the appropriateness of shared frailty models. As shown in Table 4.11 above, the likelihood ratio test statistic for the hypothesis $H_0: \rho_z = 1 \; versus \; H_1: \rho_z < 1$ is insignificant. Thus, we have no evidence to reject that the model is shared frailty model.

4.3.3 Comparison of results of semi-parametric and parametric methods

As we have seen so far, the semi-parametric shared gamma frailty model turns out to be a better fit as compared to the Cox PH model. Furthermore, the parametric shared gamma frailty model with gompertz baseline hazard was found to be appropriate than others parametric models. When we compare the results obtained from the two methods, as we can see from Table 4.6 and Table 4.9 (right), the significant covariates are the same and the heterogeneity parameter is found to be significant in both models. Although there is little difference in the magnitude of the regression coefficient estimates, the results are more or less similar.

4.3.4 Discussion of the results

This study showed that children of twin birth living in rural areas face higher risk of mortality than those living in urban areas. In agreement with this result, a child mortality study in West Africa by Balk et al (2003) found that children residing in urban areas have a better chance of survival than those residing in rural areas. Similarly, Dashtseren (2002) showed that child mortality rate is higher in rural areas than urban areas.

The current study also showed that children of twin birth who were born before the younger sibling reaches the age of 18 months experience low survival compared to those children of twin birth who have more than 24 month spacing. Several studies that are conducted on determinants of child mortality found similar results. These include studies by Becher et al (2004) in Burkina Faso and Koissi and Högnäs (2001) in Ivory Coast.

Our result revealed that first born children of twin birth experience higher risk of mortality than children of twin birth whose birth order is five and above. Similar to this result, child mortality studies by Joshua and Jeroen (2009) in Zimbabwe, Desta (2011) in Ethiopia, and Balk et al

(2003) showed that first born children experience low survival compared to those children who have higher birth order. However, the result of the study by Becher et al (2004) in Burkina Faso showed that the variable is statistically insignificant.

This study found that children of twin birth whose previous sibling died experience a lower survival chance compared to those whose previous sibling is alive or first born. A child mortality study by Koissi and Högnäs (2001) in Ivory Coast found similar results. In contrary, the variable previous sibling status was insignificant in studies by Guo and Rodríguez (1992) and Becher et al (2004).

4.4 Conclusion and recommendations

4.4.1 Conclusion

The results of this study showed that the main factors associated with twin under-five mortality are place of residence, preceding birth interval, birth order, and the status of the previous sibling. Those twins who resided in rural areas had lower chance of survival compared to those who live in urban areas. First born twins experienced high mortality compared to those twins who have birth order above five. Those twins who were born before their sibling reaches 18 month experience low survival compare to those twins who have more than 24 month spacing, and twins whose previous sibling died also experience high mortality compared to those twins whose previous sibling was alive.

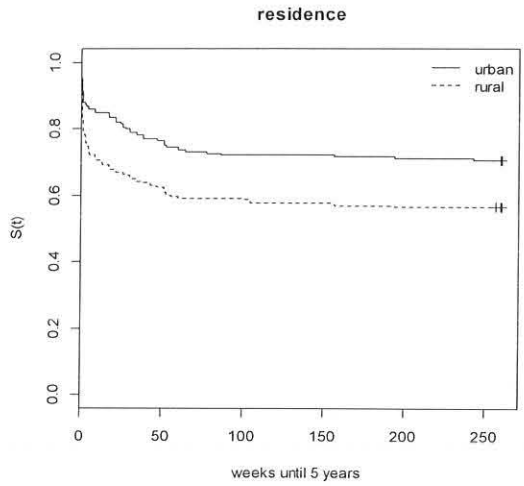
4.4.2 Recommendations

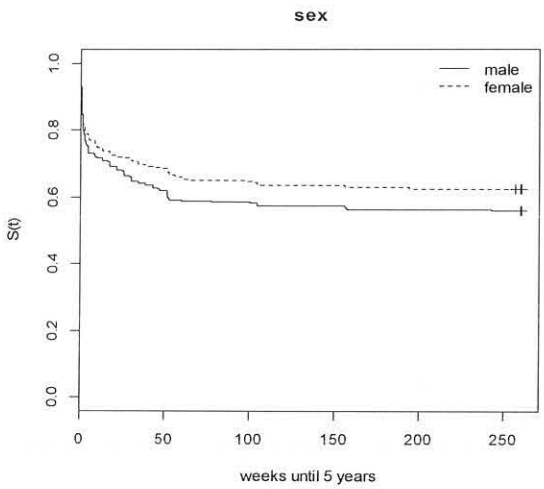
The Ethiopian government has implemented health oriented interventions to reduce child mortality. In order to reduce the rate of child mortality among twins, this study recommends the following:

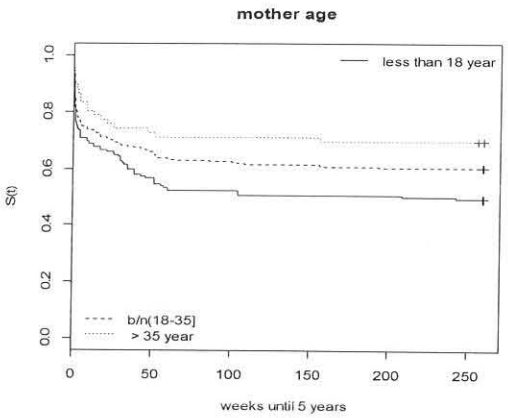
- The survival experience of children of twin birth in rural areas is much lower than in urban areas indicated that special attention should be given to improve the health infrastructure in rural areas.
- The significant effect of the birth spacing of the previous sibling on the survival chance of children of twin birth indicated that efforts have to be exerted to educate the public about family planning and birth spacing.
- The modeling practice in this study revealed the significance of including a random effect in the model. Hence, researches in this area should incorporate and assess such random effects.

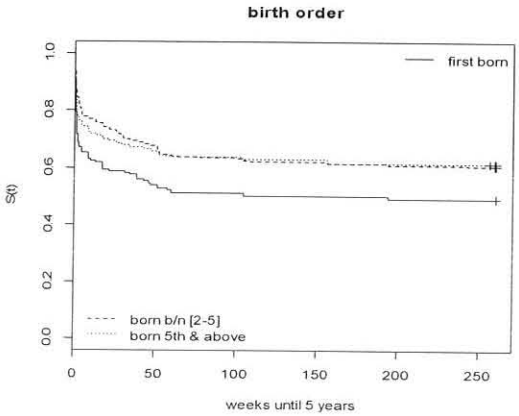
Annex

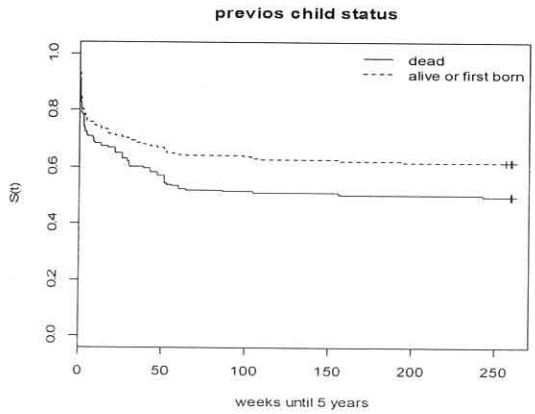
Annex A : Graphs of the estimates of the Kaplan-Meier survivor functions for each covariate.











Annex B: Likelihood function for bivariate correlated gamma frailty model with Weibull baseline hazard.

In this section for weibull baseline hazard, we drive the expression of the likelihood function for the bivariate correlated gamma frailty model with observed vectors of covariates $\mathbf{X}_1$ and $\mathbf{X}_2$. One can follow similar procedure for gompertz and exponential baseline hazard functions.

The bivariate correlated gamma frailty model is given in equation (3.46). This equation can be used to form the likelihood function of the data. However staraight forward maximization of the likelihood function require assumption about parametric form of the underlying cumulative hazard function $H_{01}(t_1)$ and $H_{02}(t_2)$. One can assume two different baseline hazard functions for the two individuals, but for twin application usually the underlying hazards functions are assumed to be identical. Suppose the underlying hazard is Weibull baseline hazard. For Weibull baseline hazard $h_0(t) = \lambda v t^{v-1}$ (where $\lambda>0$ and $v > 0$ are the scale and shape parameters respectively), the cumulative baseline hazard is given by: $H_0(t) = \lambda t^v$. Thus the bivariate correlated gamma frailty model (3.46) with observed vectors of covariates X_1 and $X_2$is given by:

$$S(t_1, t_2 \mid X_1, X_2) = \left(1 + \sigma_z^2 \lambda {t_1}^v e^{\beta' X_1} + \sigma_z^2 \lambda {t_2}^v e^{\beta' X_2}\right)^{-\frac{\rho_z}{\sigma_z^2}} \times \left(1 + \sigma_z^2 \lambda {t_1}^v e^{\beta' X_1}\right)^{\frac{-1+\rho_z}{\sigma_z^2}} \left(1 + \sigma_z^2 \lambda {t_2}^v e^{\beta' X_2}\right)^{\frac{-1+\rho_z}{\sigma_z^2}} \quad (B1)$$

This expression (B1) can be used to form the likelihood function for the bivariate correlated gamma frailty model given in (3.51). The likelihood function can be obtained by using equation (3.16). As a result, the expression $S_{t_{i1},t_{i2}}(t_{i1}, t_{i2})$ in (3.51) can be obtained by taking the differentiation of (B1) with respect to t_1, t_2 . Thus for Weibull baseline hazard and with observed vectors of covariates $X_1$and X_2 this expression is given by:

$$S_{t_{i1},t_{i2}}(t_{i1}, t_{i2} \mid X_1, X_2) = \left[(\lambda v {t_1}^{v-1} e^{\beta' X_1})(\lambda v {t_2}^{v-1} e^{\beta' X_2}) \times (n)^{-\frac{(1-\rho_z)}{\sigma_z^2}} \times (p)^{-\frac{(1-\rho_z)}{\sigma_z^2}} (m)^{-\frac{\rho_z}{\sigma_z^2}}\right] \times [A]$$

where $A=\left[\frac{(1-\rho_z)\rho_z}{(m)} \times \left(\frac{1}{(n)} + \frac{1}{(p)}\right) + \frac{(\rho_z+\sigma_z^2)\rho_z}{(m)^2} + \frac{(1-\rho_z)^2}{(n)(p)}\right]$

$$m = n + p - 1$$

$$n = 1 + \sigma_z^2 \lambda {t_1}^v e^{\beta' X_1}$$

$$p = 1 + \sigma_z^2 \lambda {t_2}^v e^{\beta' X_2}$$

The expression $S_{t_{i1}}(t_{i1}, t_{i2})$ in (3.51) can be obtained by taking the partial differentiation of (B1) with respect to t_1 . Thus for weibull baseline hazard and with observed vectors of covariates X_1 and X_2 this expression is given by:

$$S_{t_{i1}}(t_{i1}, t_{i2} \mid X_1, X_2) = \left(\lambda v {t_1}^{v-1} e^{\beta \cdot X_1}\right)\left(1 + \sigma_z^2 \lambda {t_1}^{v} e^{\beta \cdot X_1}\right)^{-\frac{(1-\rho_z)}{\sigma_z^2}-1}$$
$$\times \left(1 + \sigma_z^2 \lambda {t_2}^{v} e^{\beta \cdot X_2}\right)^{-\frac{(1-\rho_z)}{\sigma_z^2}} \times 1 + \sigma_z^2 (\lambda {t_2}^{v} e^{\beta \cdot X_2} + \lambda {t_1}^{v} e^{\beta \cdot X_1})^{-\frac{\rho_z}{\sigma_z^2}-1}$$
$$\times \left(\left(1 + \sigma_z^2 \lambda {t_1}^{v} e^{\beta \cdot X_1}\right) + \left((1 - \boldsymbol{\rho_z})\right)\left(1 + \sigma_z^2 \lambda {t_2}^{v} e^{\beta \cdot X_2}\right) - \left((1 - \boldsymbol{\rho_z})\right)\right)$$

The expression $S_{t_{i2}}(t_{i1}, t_{i2})$ in (3.51) can be obtained by taking the partial differentiation of (B1) with respect to t_2 . Thus for WWeibull baseline hazard and with observed vectors of covariates $X_1$and X_2 this expression is given by:

$$S_{t_{i2}}(t_{i1}, t_{i2} \mid X_1, X_2) = \left(\lambda v {t_2}^{v-1} e^{\beta \cdot X_2}\right)\left(1 + \sigma_z^2 \lambda {t_2}^{v} e^{\beta \cdot X_2}\right)^{-\frac{(1-\rho_z)}{\sigma_z^2}-1}$$
$$\times \left(1 + \sigma_z^2 (\lambda {t_2}^{v} e^{\beta \cdot X_2} + \lambda {t_1}^{v} e^{\beta \cdot X_1})\right)^{-\frac{\rho_z}{\sigma_z^2}-1} \times \left(1 + \sigma_z^2 \lambda {t_1}^{v} e^{\beta \cdot X_1}\right)^{-\frac{(1-\rho_z)}{\sigma_z^2}}$$
$$\times \left(\left(1 + \sigma_z^2 \lambda {t_2}^{v} e^{\beta \cdot X_2}\right) + \left((1 - \boldsymbol{\rho_z})\right)\left(1 + \sigma_z^2 \lambda {t_1}^{v} e^{\beta \cdot X_1}\right) - \left((1 - \boldsymbol{\rho_z})\right)\right)$$

The expression $S(t_{i1}, t_{i2})$ in (3.51) is the survival function of the bivariate correlated gamma frailty model. Thus for Weibull baseline hazard and with observed vectors of covariates $X_1$and X_2 this expression is therefore the expression given in (B1) above.

Plugging these four expressions into equation (3.51) gives the likelihood function of the bivariate correlated gamma frailty model with observed covariates. Usually the logarithm of the likelihood function is used in parameter estimation. Now the estimates of the parameters $(\sigma_z^2, \rho_z, \beta, \lambda, v)$ can be obtained using the standard maximum likelihood parameter estimation procedure.

Annex C: Procedure of simulation and results

1. Procedure of simulation

The bivariate parametric PH models that are used in this study are fitted using self written program on R software. The programs are tested on simulation data sets. Here we describe the procedures of simulation of the survival times in a bivariate parametric PH models.

Suppose $S(t \mid X)$ is the survival function of the Cox proportional hazards model and is given by:

$$S(t \mid X) = e^{-e^{\beta' X} \mathrm{H_0(t)}} \qquad (c1)$$

The distribution function of the above Cox model (c1) is:

$$F(t|\mathrm{x}) = 1 - e^{-e^{\beta' X} \mathrm{H_0(t)}} \qquad (c2)$$

Let y be a random variables with distribution function $F(.)$, then $u = F(y)$ follows a uniform distribution on the interval from 0 to 1. Furthermore, if $u \sim U[0,1]$, then 1-$u \sim U[0,1]$ (Bender et al, 2005). Thus, let T be the survival times of model (c1), then it follows from (c2) that:

$$u = \exp\left(-e^{\beta' X} \mathrm{H_0(t)}\right) \sim U[0,1]$$

If the baseline hazard function $h_0(t) > 0$ for all t, then $\mathrm{H_0(t)}$ can be inverted and the survival time T can be expressed as:

$$T = H_O^{-1}\left[-\log(\mathrm{u}) e^{\beta' X}\right] \qquad (c3)$$

The conditional survival function in the bivariate correlated gamma frailty model with observed vectors of covariates $\mathrm{X_1}$ and $\mathrm{X_2}$ is given by:

$$S(t_1, t_2|Z_1, Z_2) = S_1(t_1|Z_1) S_2(t_2|Z_2) = e^{-Z_1 e^{\beta' X_1} H_{01}(t_1)} e^{-Z_2 e^{\beta' X_2} H_{02}(t_2)} \qquad (c4)$$

where Z_1 and Z_2 are correlated gamma-distributed frailty variables, $S(t \mid Z)$ denotes the survival function of an individual conditional on the frailty variable Z. Usually this model assumes proportional hazard structure. Which means the proportional hazard model is obtained if the frailty distribution degenerates to Z=1.

Suppose there are n pairs. Conditional on the frailty terms Z_{i1} and Z_{i2} , the survival times T_{i1} *and* T_{i2} in a pair i are assumed to be independent. Hence by including the frailties variables as additional covariates with regression coefficient of one, we can generate the survival times

T_{i1} *and* T_{i2} . If $Z_{i1} = Z_{i2} = 1$ $(i = 1,2,\dots,n)$ the bivariate survival function of the bivariate correlated gamma frailty model (c4) is given by:

$$S(t_1, t_2) = S_1(t_1)S_2(t_2) = e^{-e^{\beta' X_1} H_{01}(t_1)} e^{-e^{\beta' X_2} H_{02}(t_2)} \qquad \text{(c5)}$$

This means a bivariate PH model without frailty. Thus the survival times T_{i1} *and* T_{i2} can be generated by using (c3).

If $Z_{i1} = Z_{i2} = Z_i$ $(i = 1,2,\dots,n)$ the bivariate survival function of the bivariate correlated gamma frailty model (c4) is given by:

$$S(t_1, t_2|Z) = S_1(t_1|Z)S_2(t_2|Z) = e^{-Ze^{\beta' X_1} H_{01}(t_1)} e^{-Ze^{\beta' X_2} H_{02}(t_2)} \qquad (C6)$$

This is the shared gamma frailty model. Given the common frailty Z, the survival times are independent. Hence the survival times T_{i1} *and* T_{i2} can be generated by:

$$T_{ij} = H_0^{-1}\left[-\log(u_{ij})\, Z e^{\beta' X_{ij}}\right] \quad for\ j = 1,2 \qquad \text{(c7)}$$

Let Z_1 and Z_2 be given by (3.43) which is the correlated gamma frailty model. Given the frailties Z_{i1} and Z_{i2} , the survival times are independent. Hence the survival times T_{i1} *and* T_{i2} can be generated by:

$$T_{ij} = H_0^{-1}\left[-\log(u_{ij}) Z_{ij} e^{\beta' X_i}\right] \quad for\ j = 1,2 \qquad \text{(c8)}$$

The characteristics for a Cox model with weibull and gompertz baseline hazard are described as follows:

Characteristics	Weibull model	Gompertz model
parameters	$\lambda>0$ is the scale $v > 0$ is the shape	$\lambda>0$ is the scale $\varphi\ (-\infty, \infty)$ is the shape
Baseline hazard function	vt^{v-1}	$\lambda \exp(\varphi t)$
Cumulative baseline hazard function	λt^{v}	$\frac{\lambda}{\varphi}\ (exp(\varphi t) - 1)$
Inverse cumulative hazard function	$\left(\frac{1}{\lambda} t\right)^{1/v}$	$\frac{1}{\varphi} log\left[\frac{\varphi}{\lambda} t + 1\right]$
Survival time using	$T_i = \left(-\frac{\log(u_i)}{\lambda Z_i e^{\beta' X_i}}\right)^{1/v}$	$T_i = \frac{1}{\varphi} log\left[1 - \frac{\varphi \log(u_i)}{\lambda Z_i e^{\beta' X_i}}\right]$

2. Results of the simulation study

2.1 Results of the Bivariate PH models without frailty

a. Weibull baseline hazard with parameters $= 2$ and $\lambda = 0.002$. 100 simulated dataset were generated. Each artificial dataset contained 1000 pairs of independent weibull distributed durations T_1 and T_2 . One continuous (U [0,1]) and one categorical (Bernoulli (0.7)) with regression coefficients $\beta_1 = 2$ and $\beta_2 = -1$ respectively were used as observed covariates. Each observed covariates for the two individuals in a pair were taken to be independent. 25% of the events were censored. The results of the bivariate weibull PH model without frailty and the results obtained from the generic R function *coxph* are presented below.

MOE=Mean of Estimates of 100 datasets

MoSE= Mean of Standard Errors of 100 datasets

Stand dev=Standard deviation of the 100 estimates

	Results of Weibull model without frailty			Results from coxph	
parameters	MOE	MoSE	Stand dev	MOE	MoSE
$\beta_1 = 2$	2.0013	0.08245	0.0831	2.0048	0.0961
$\beta_2 = -1$	-1.0024	0.05542	0.0565	-1.0091	0.0556
$v = 2$	2.0078	0.02285	0.0237	----	------
$\lambda = 0.002$	0.002009	0.000096	0.00010	---	-----

b. Gompertz baseline hazard with parameters $\varphi = 0.1$ and $\lambda = 0.01$. 100 simulated dataset were generated. Each artificial dataset contained 1000 pairs of independent Gompertz distributed durations T_1 and T_2 . One continuous (U [0,1]) and one categorical (Bernoulli (0.7)) with regression coefficients $\beta_1 = 2$ and $\beta_2 = -1$ respectively were used as observed covariates. Each observed covariates for the two individuals in a pair were taken to be independent. 25% of the events were censored. The results of the bivariate weibull PH model without frailty and the results obtained from the generic R function *coxph* are presented below.

	Results of gompertz PH model without frailty			Results from coxph	
parameters	MOE	MoSE	Stand dev	MOE	MoSE
$\beta_1 = 2$	2.0013	0.08245	0.0831	2.0048	0.0961
$\beta_2 = -1$	-1.0024	0.05542	0.0565	-1.0091	0.0556
$\varphi = 0.1$	2.0078	0.02285	0.0237	----	------
$\lambda = 0.01$	0.002009	0.000096	0.00010	---	-----

2.2 Results of the Bivariate PH models with Shared gamma frailty model

For each of the two baseline hazard 100 data sets were generated , each data set have 1000 pairs of related durations. For weibull the shared frailty parameter was $\sigma_z^2 = 0.5$ and for gompertz the shared frailty parameter was $\sigma_z^2 = 0.25$. One continuous (U [0,1]) and one categorical (Bernoulli (0.7)) with regression coefficients $\beta_1 = 3$ and $\beta_2 = -1$ respectively were used as observed covariates. Each observed covariates for the two individuals in a pair were taken to be independent and 25% of the events are censored.

	Results of Weibull model with shared gamma frailty			Results from coxph with shared gamma frailty	
parameters	MOE	MoSE	Stand dev	MOE	MoSE
$\beta_1 = 3$	2.9993	0.09748	0.09989	2.9999	0.12225
$\beta_2 = -1$	-1.0040	0.07244	0.07567	-1.0068	0.07083
$\sigma_z^2 = 0.5$	0.5002	0.05652	0.05918	0.4997	----
$v = 2.5$	2.50001	0.031175	0.0237	----	------
$\lambda = 0.01$	0.00100	0.000022	0.000034	---	-----

	Results of gompertz model with shared gamma frailty			Results from coxph with shared gamma frailty	
parameters	MOE	MoSE	Stand dev	MOE	MoSE
$\beta_1 = 3$	3.0012	0.1332	0.14334	3.0004	0.11943
$\beta_2 = -1$	-1.0054	0.0689	0.0697	-1.0057	0.0661
$\sigma_z^2 = 0.25$	0.25208	0.05411	0.05710	0.256	----
$\varphi = 0.2$	0.20221	0.00118	0.00112	----	------
$\lambda = 0.02$	0.00202	0.00209	0.002001	---	-----

2.3 Correlated gamma frailty model

a. Gompertz baseline hazard with parameters $\varphi = 0.1$ *and* $\lambda = 0.01$. 1000 data sets were generated, each data set have 2000 pairs of related individuals. The frailty parameters are taken to be $\sigma_z^2 = 0.25$ *and* $\rho_z = 0.5$. One continuous (U [0,1]) and one categorical (Binomial (2,0.3)) with regression coefficients $\beta_1 = 2$ and $\beta_2 = -1$ respectively were used as observed covariates. Each observed covariates for the two individuals in a pair were taken to be independent and 25% of the events are censored.

parameters	True value	MOE	MoSE	Stand dev
σ_z^2	0.25	0.248102	0.061137	0.059744
ρ_z	0.5	0.5129727	0.1832322	0.1798843
φ	0.02	0.02001794	0.0018299	0.0018194
λ	0.01	0.01001131	0.0006103	0.00061501
β_1	2	1.997301	0.0988037	0.09622935
β_2	-1	-1.001258	0.0484540	0.04722147

b. weibull baseline hazard with parameters $v = 2.5$ and $\lambda = 0.001$. 1000 data sets were generated, each data set have 2000 pairs of related individuals. The frailty parameters are taken to be $\sigma_z^2 = 0.5$ and $\rho_z = 0.6$. One continuous (U [0,5]) and one categorical (Binomial (2,0.3)) with regression coefficients $\beta_1 = 1$ and $\beta_2 = -1$ respectively were used as observed covariates. Each observed covariates for the two individuals in a pair were taken to be independent and 25% of the events are censored.

parameters	True value	MOE	MoSE	Stand dev
σ_z^2	0.5	0.5020934	0.05383121	0.0570779
ρ_z	0.6	0.6003374	0.08389226	0.08610009
v	2.5	2.504573	0.06052938	0.063607
λ	0.001	0.00100013	0.00015211	0.00016285
β_1	1	1.002376	0.0291718	0.03074984
β_2	-1	-1.001968	0.04474191	0.04673245

Annex D: R codes for bivariate gamma correlated gamma frailty with two observed covariates: Weibull baseline hazard

```
x=matrix(c(t1,t2,zx1,zx2,zy1,zy2,cen1,cen2),2000,8)    # x is matrix of survival times t1 and t2,
#covariates (zx and zy ) and censor statuses  cen1 and cen2.
thet=c(0.01,0.02,0.000009,0.08,1.5 ,0.4 )      # initial parameters or
# parameters (variance b, correlation R, weibull  parameters (a1,a2)  and regression coefficients
#(c1and c2)

loglikweib = function (thet, x)
{   x=as.matrix(x); t1=x[,1]; t2=x[,2]  ;y11=x[,3]; y12=x[,4]; y21=x[,5]; y22=x[,6]  ;cen1=x[,7];
cen2=x[,8]; b = thet[1]; R = thet[2]; a1 = thet[3]  ;a2 = thet[4];c1 = thet[5] ; c2 = thet[6]
k1=c1*y11+c2*y21; k2=c1*y12+c2*y22
n= (b*a1*exp(k1)*t1^a2+1); p=(b*a1*exp(k2)*t2^a2+1)
m= ((b*a1*exp(k1)*t1^a2+1)+(b*a1*exp(k2)*t2^a2))
A=((b+R)*R/(m)^2)+((R*(1-R))/(m))*(1/n+1/p)+((1-R)^2/(n*p))

# the log likelihood function
l=sum(cen1*cen2*(2*log(a1)+2*log(a2)+(a2-1)*(log(t1)+log(t2))+k1+k2-((1-
R)/b)*(log(n)+log(p))-(R/b)*log(m)+log(A))+
cen1*(1-cen2)*(log(a1)+log(a2)+(a2-1)*log(t1)+k1-((1-R)/b+1)*log(n)-((1-R)/b)*log(p)-
(R/b+1)*log(m)+log((n+(1-R)*p-(1-R))))+
cen2*(1-cen1)*(log(a1)+log(a2)+(a2-1)*log(t2)+k2-((1-R)/b+1)*log(p)-((1-R)/b)*log(n)-
(R/b+1)*log(m)+log((p+(1-R)*n-(1-R))))+
(1-cen2)*(1-cen1)*(-((1-R)/b)*(log(n)+log(p))-R/b*log(m)))
return(l)
}

#gradient function

gradweib=function(thet,x){
x=as.matrix(x); t1=x[,1]; t2=x[,2]  ;y11=x[,3]; y12=x[,4]; y21=x[,5]; y22=x[,6]  ;cen1=x[,7];
cen2=x[,8]; b = thet[1]; R = thet[2]; a1 = thet[3]  ;a2 = thet[4];c1 = thet[5] ; c2 = thet[6]
k1=c1*y11+c2*y21; k2=c1*y12+c2*y22
n=(b*a1*exp(k1)*t1^a2+1); p=(b*a1*exp(k2)*t2^a2+1)
m=((b*a1*exp(k1)*t1^a2+1)+(b*a1*exp(k2)*t2^a2))
A=((b+R)*R/(m)^2)+((R*(1-R))/(m))*(1/n+1/p)+((1-R)^2/(n*p))

#first derivatives of n, p, m and A w.r.t the parameters
n1=(a1*exp(k1)*t1^a2);  p1=(a1*exp(k2)*t2^a2);  m1=n1+p1
n3=(b*exp(k1)*t1^a2);  p3=(b*exp(k2)*t2^a2);  m3=(n3+p3)
n4=(b*a1*exp(k1)*t1^a2*log(t1));  p4=(b*a1*exp(k2)*t2^a2*log(t2));  m4=n4+p4
n5=y11*b*a1*exp(k1)*t1^a2;  p5=y12*b*a1*exp(k2)*t2^a2;  m5=n5+p5
n6=y21*b*a1*exp(k1)*t1^a2;  p6=y22*b*a1*exp(k2)*t2^a2;  m6=n6+p6
A1=((R/m^2)-2*R*(R+b)*(1/m^3)*m1-(R-R^2)*(1/m^2)*m1*(1/n+1/p)+(-(1/n^2)*n1-
(1/p^2)*p1)*((R-R^2)/m)+(1-R)^2*(-(1/((n*p)^2))*(n1*p+p1*n)))
A2=((2*R+b)*(1/(m^2))+(1-2*R)*(1/m)*(1/n+1/p)+(-2+2*R)*(1/(n*p)))
```

```
A3=(-(2*R*(R+b))*(1/(m^3))*m3-(R-R^2)*(1/(m^2))*m3*((1/n)+(1/p))+(-(1/(n^2))*n3-
(1/(p^2))*p3)*((R-R^2)/m)+(1-R)^2*(-(1/((n*p)^2))*(n3*p+p3*n)))
A4=(-(2*R*(R+b))*(1/(m^3))*m4-(R-R^2)*(1/(m^2))*m4*((1/n)+(1/p))+(-(1/(n^2))*n4-
(1/(p^2))*p4)*((R-R^2)/m)+(1-R)^2*(-(1/((n*p)^2))*(n4*p+p4*n)))
A5=(-(2*R*(R+b))*(1/(m^3))*m5-(R-R^2)*(1/(m^2))*m5*((1/n)+(1/p))+(-(1/(n^2))*n5-
(1/(p^2))*p5)*((R-R^2)/m)+(1-R)^2*(-(1/((n*p)^2))*(n5*p+p5*n)))
A6=(-(2*R*(R+b))*(1/(m^3))*m6-(R-R^2)*(1/(m^2))*m6*((1/n)+(1/p))+(-(1/(n^2))*n6-
(1/(p^2))*p6)*((R-R^2)/m)+(1-R)^2*(-(1/((n*p)^2))*(n6*p+p6*n)))
```

#the gradient matrix

```
g=c(sum(cen1*cen2*(((1-R)/b^2)*(log(n)+log(p))-((1-R)/b)*((1/n)*n1+(1/p)*p1)+
(R/b^2)*log(m)-(R/b)*((1/m)*m1)+(1/A)*A1)+ cen1*(1-cen2)*(((1-R)/b^2)*(log(n)+log(p))-
((1-R)/b+1)*(1/n)*n1-((1-R)/b)*(1/p)*p1+(R/b^2)*log(m)-(R/b+1)*(1/m)*m1+
(1/(n+(1-R)*p-(1-R)))*(n1+(1-R)*p1))+ cen2*(1-cen1)*(((1-R)/b^2)*(log(n)+log(p))-
((1-R)/b+1)*(1/p)*p1-((1-R)/b)*(1/n)*n1+(R/b^2)*log(m)-(R/b+1)*(1/m)*m1+
(1/(p+(1-R)*n-(1-R)))*(p1+(1-R)*n1))+ (1-cen2)*(1-cen1)*(((1-R)/b^2)*(log(n)+log(p))-
((1-R)/b)*((1/n)*n1+(1/p)*p1)+(R/b^2)*log(m)-(R/b)*(1/m)*m1)),
sum(cen1*cen2*(1/b*(log(n)+log(p))-1/b*log(m)+(1/A)*A2)+cen1*(1-cen2)*(1/b*(log(n)+log(p))-
1/b*log(m)+(1/(n+(1-R)*p-(1-R)))*(1-p))+ cen2*(1-cen1)*(1/b*(log(n)+log(p))-1/b*log(m)+
(1/(p+(1-R)*n-(1-R)))*(1-n))+(1-cen2)*(1-cen1)*((1/b)*(log(n)+log(p))-1/b*log(m))),
sum(cen1*cen2*(2/a1-((1-R)/b)*((1/n)*n3+(1/p)*p3)-(R/b)*((1/m)*m3)+(1/A)*A3)+
cen1*(1-cen2)*(1/a1-((1-R)/b+1)*(1/n)*n3-((1-R)/b)*(1/p)*p3-(R/b+1)*(1/m)*m3+
(1/(n+(1-R)*p-(1-R)))*(n3+(1-R)*p3))+ cen2*(1-cen1)*(1/a1-((1-R)/b+1)*(1/p)*p3-
((1-R)/b)*(1/n)*n3-(R/b+1)*(1/m)*m3+(1/(p+(1-R)*n-(1-R)))*(p3+(1-R)*n3))+
(1-cen2)*(1-cen1)*(-((1-R)/b)*((1/n)*n3+(1/p)*p3)-(R/b)*(1/m)*m3)),
sum(cen1*cen2*(2/a2+(log(t1)+log(t2))-((1-R)/b)*((1/n)*n4+(1/p)*p4)-(R/b)*((1/m)*m4)+(1/A)*A4)+
cen1*(1-cen2)*(1/a2+log(t1)-((1-R)/b+1)*(1/n)*n4-((1-R)/b)*(1/p)*p4-(R/b+1)*(1/m)*m4+
(1/(n+(1-R)*p-(1-R)))*(n4+(1-R)*p4))+cen2*(1-cen1)*(1/a2+log(t2)-((1-R)/b+1)*(1/p)*p4-
((1-R)/b)*(1/n)*n4-(R/b+1)*(1/m)*m4+(1/(p+(1-R)*n-(1-R)))*(p4+(1-R)*n4))+
(1-cen2)*(1-cen1)*(-((1-R)/b)*((1/n)*n4+(1/p)*p4)-(R/b)*(1/m)*m4)),
sum(cen1*cen2*((y11+y12)-((1-R)/b)*((1/n)*n5+(1/p)*p5)-(R/b)*((1/m)*m5)+(1/A)*A5)+
cen1*(1-cen2)*(y11-((1-R)/b+1)*(1/n)*n5-((1-R)/b)*(1/p)*p5-(R/b+1)*(1/m)*m5+
(1/(n+(1-R)*p-(1-R)))*(n5+(1-R)*p5))+ cen2*(1-cen1)*(y12-((1-R)/b+1)*(1/p)*p5-
((1-R)/b)*(1/n)*n5-(R/b+1)*(1/m)*m5+(1/(p+(1-R)*n-(1-R)))*(p5+(1-R)*n5))+
(1-cen2)*(1-cen1)*(-((1-R)/b)*((1/n)*n5+(1/p)*p5)-(R/b)*(1/m)*m5)),
sum(cen1*cen2*((y21+y22)-((1-R)/b)*((1/n)*n6+(1/p)*p6)-(R/b)*((1/m)*m6)+(1/A)*A6)+
cen1*(1-cen2)*(y21-((1-R)/b+1)*(1/n)*n6-((1-R)/b)*(1/p)*p6-(R/b+1)*(1/m)*m6+
(1/(n+(1-R)*p-(1-R)))*(n6+(1-R)*p6))+ cen2*(1-cen1)*(y22-((1-R)/b+1)*(1/p)*p6-
((1-R)/b)*(1/n)*n6-(R/b+1)*(1/m)*m6+(1/(p+(1-R)*n-(1-R)))*(p6+(1-R)*n6))+
(1-cen2)*(1-cen1)*(-((1-R)/b)*((1/n)*n6+(1/p)*p6)-(R/b)*(1/m)*m6)))
return(g)}
```

```
# constrOptim for the natural constraints such as variance >0 and corelation [0,1].
# to obtain the estimates
fit=constrOptim(thet, loglikweib,x=x,grad=gradweib ,
ui=matrix(c(1,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0), nrow=3), ci=c(0,0,-1.000001), hessian=T
,control=list(fnscale=-1))
fit
```

References

1. Aaby, P. Pison, G. and Andersen, M (1995). Lower mortality for female-female twins than male-male and male-female twins in rural Senegal . US National Library of Medicine National institutes of Health , 6(4):419-422 .
2. Alam N, Van Ginneken JK, Bosch AM.(2007). Infant mortality among twins and triplets in rural Bangladesh in 1975–2002. Trop Med Int Health, **12**(12)**:**1506-1514.
3. Andersen, P.K., Klein, J.P., Knudsen, K., and Palacios, R.T (1995). Estimation of variance in Cox's regression model with gamma frailties. Technical report, Department of Biostatistics, University of Copenhagen, Denmark.
4. Andersen, P. K., Borgan, O., Gaill, R. D., and Keiding, N. (1993). Statistical Methods Based on Counting Processes. Springer, New York.
5. Asefa, M., Drewett, R., Tessema, F. (2000). A birth cohort study in south-west Ethiopia to identify factors associated with infant mortality that are amenable for intervention. Ethiopian Journal of Health Development. 14(2). Pages 161-168.
6. Balk Deborah, Tom Pullum, Adam Storeygard, Fern Greenwell, and Melissa Neuman (2003) Spatial Analysis of Childhood Mortality in West Africa. Calverton, Maryland, USA: ORC Macro and Center for International Earth Science Information Network (CIESIN), Columbia University.
7. Beard, R.E. (1959). Note on some mathematical mortality models. In: The Lifespan of Animals. Ciba Foundation Colloquium on Ageing, Little, Brown, Boston, 302–311.
8. Becher, H., Muller, O., Jahn, A., Gbangou, A., Kynast-Wolf, G., and Kouyate, B. (2004). Risk factors of infant and child mortality in rural Burkina Faso. Bulletin of the World Health Organization, 82:265-273.
9. Bender, R., Augustin, T., and Blettner, M (2005). Generating survival times to simulate Cox proportional hazards models. Statist. Med, 24:1713–1723.
10. Breslow, N.(1974). Covariance analysis of censored survival data.Biometrics, 30: 89–99.
11. Bryce J, Black RE, Walker N, Bhutta ZA, Lawn JE, Steketee RW. (2005). Can the world afford to save the lives of 6 million children each year?. Lancet, **365**(9478)**:**2193-2200.
12. Boyles, R. A. (1983). On the convergence of the EM algorithm. Journal of the Royal Statistical Society Series B 45: 47–50.
13. Bulmer, M. G. (1980). The mathematical theory of quantitative genetics. Clarendon Press, Oxford.
14. Childs, B. Moxon, R. and Winkelstein, J. (1992). Genetics and Infectious Diseases. The Genetic Basis of Common Diseases. Oxford University Press.
15. Christensen, K., Vaupel, J., Holm, N., Yashin, A. (1995). Mortality among twins after age 6: fetal origins hypothesis versus twin method. British Medical Journal 310, 432-436.
16. Claeskens, G., Nguti, R., and Janssen, p (2008). One-sided tests in shared frailty models. OR & Business Statistics.

17. Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65 (1): 141-151.

18. Congdon, P. (1995). Modelling frailty in area mortality. Statistics in Medicine,14: 1859–74.

19. Cox, D.R. (1972). Regression models and life tables. Journal of the Royal Statistical Society, B, 34:187-220.

20. Cox, D.R. (1975). Partial likelihood. Biometrika, 62, 269-76

21. Cox, D. R. and Oakes, D. Analysis of Survival Data. (1984) Chapman and Hall, New York.

22. Cox, D. R., and Snell, E. J. (1968). A general definition of residuals with discussion. Journal of the Royal Statistical Society, Series B 30, 248.275.

23. Dashtseren, A. (2002). Determinants of infant and child mortality in Mongolia. Paper presented at the IUSSP Regional Conference, Bangkok, Thailand.

24. Dayton, C.M (2003). Model Comparisons Using Information Measures. Journal of Modern Applied Statistical Methods, 2: 281-292.

25. Dempster, A.P., Laird, N.M., Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society (B) 39: 1–38.

26. Deribew, A., Tessema, F. and Girma, B (2007). Determinants of under-five mortality in Gilgel Gibe Field Research Center, Southwest Ethiopia. Ethiopian Journal of Health Development. 21(2). Pages 1-8.

27. Desta Mekonnen (2011). Infant and Child Mortality in Ethiopia: The role of socioeconomic, demographic and biological factors in the previous five years period of 2000 and 2005. Lund University.

28. DHS. The 2011 Ethiopian Demographic and Health Surveys. Measure DHS. 2011; Available from: www.measuredhs.com.

29. dos Santos, D., Davies, R, Francis, B. (1995). Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. Journal of Statistical Planning and Inference, 47: 111–127.

30. Duchateau, L., Janssen, P. The Frailty Model. (2008) Springer, New York.

31. Duffie, Darrell, Andreas Eckner, Guillaume Horel, and Leandro Saita, (2009). Internet Appendix to "Frailty Correlated Default'', Journal of Finance. 64, 2089-2123, http://www.afajof.org/IA/2009.asp.

32. Efron, B and Hinkley, D.V. (1978). Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. Biometrika, 65(3):457-87.

33. Falconer, D. S. Introduction to Quantitative Genetics. (1990) Longman Scientific and Technical, NewYork

34. Fottrell, E., Enquselassie, E., and Byass, P. The distribution and effects of child mortality risk factors in Ethiopia: A comparison of estimates from DSS and DHS. (2009), Ethiopian Journal of Health Development, 23(2): 163-168.

35. Garibotti, G, Smith, K.R., Kerber, R.A. and Boucher, K.M (2006). Longevity and Correlated Frailty in Multigenerational Families. The Gerontological Society of America, 12, 1253–1261.

36. Giard N (2001). A four-dimensional correlatedfrailty model. Ph.D thesis, Faculty of Mathematics and Natural Sciences, University of Rostock.

37. Greenwood, M., Yule, G.U. (1920) . An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. Journal of the Royal Statistical Society, 83: 255–279.

38. Guo G. and Rodriguez G. (1992). Estimating a Multivariate Proportional Hazards Model for Clustered Data Using the EM Algorithm, with an Application to Child Survival in Guatemala. Journal of American Statistical Association. 87: 969-976.

39. Guo, G., and Grummer-Strawn, L., (1993). Child Mortality Among Twins in Less Developed Countries. Population Studies: A Journal of Demography , 47:495-510.

40. Hanagal, D.D. Modeling survival Data Using Frailty Models. (2010) Taylor and Francis Group.

41. Hosmer D.W, Lemeshow S. and May S. (2008) Applied survival analysis regression modeling of time to event data. John Wiley and Sons, Inc., New York.

42. Hougaard, P. Analysis of Multivariate Survival Data. (2000). Springer: New York.

43. Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. Biometrika, 73: 387–396.

44. Hougaard, P., Harvald, B., and Holm, N. V. (1992). Measuring similarities between the lifetimes of adult Danish twins born between 1881-1830. *JASA* 87: N 417.

45. Iachine, I.A. (1995). Correlated frailty concept in the analysis of bivariate survival data. Bachelor project, Department of Mathematics and Computer Science, Odense University, Denmark.

46. Iachine, I.A. (2006). Identifiability of bivariate frailty models. Working Paper, Department of Statistics and Demography, University of Southern Denmark, Denmark.

47. Iachine, I.A. (2002). The use of twin and family survival data in the population studies of aging: Statistical methods based on multivariate survival data. Ph.D. thesis, Department of Statistics and Demography, University of Southern Denmark, Denmark.

48. Iachine, I.A., Holm N., Harris, J.R., Skythe, A. Begun, A.Z., Iachina, M.K., Lai tinen, M., Kaprio, J., and Yashin, A.I. (1998). How heritable is individual susceptibility to death? The resul ts of an analysis of survival data on Danish, Swedish and Finnish twins. Twin Research, 1, 196–205.

49. Jiezhi Qi. (2009). Comparison of Proportional Hazards and Accelerated Failure Time Models. MSc Thesis, Department of Mathematics and Statistics, University of Saskatchewan, Saskatchewan.

50. Joshua K. and Jeroen G. (2009). Determinants of infant and child mortality in Zimbabwe: Result of multivariate hazard analysis. Demographic Research, 21: 367-384.

51. Justesen, A., and Kunst, A. (2000). Postneonatal and child mortality among twins in Southern and Eastern Africa. International Journal of Epidemiology, 29:678–683.

52. Kalbfleisch, J.D., and Prentice, R. L. (1973). Marginal likelihoods based on Cox's regression and life model. Biometrika, 60:267-278.

53. Kalbfleisch, J.D and Prentice, R. The Statistical Analysis of Failure Time Data. (2002) John Wiley & Sons: New York.

54. Kalbfleish, J. D., and Prentice, R. L. The Statistical Analysis of Failure Type Data. (1980) New York,Wiley.

55. Kaplan, E. and Meier, P.(1958). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53, 457–481.

56. Klein, J.P. (1992). Semi parametric estimation of random effects using the Cox model based on EM algorithm. Biometrics, 48, 795-806.

57. Klein, J.P. and Moeschberger, M.L. (2003). Survival Analysis. Second edition. Springer: New York.

58. Klein, J. P., and Moeschberger, M. L. (1997) Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York.

59. Koissi M.C. and Högnäs, G. (2001). Using WinBUGS to Study Family Frailty in Child Mortality, with an Application to Child Survival in Ivory Coast. African Population Studies, Vol.20 n°1.

60. Lancaster, T. (1979). Econometric methods for the duration of unemployment. Econometrica, 47:939–956.

61. Lee, E. T., and Wang, J. W. Statistical Methods for Survival Data Analysis, 3rd ed. Wiley, New York, 2003.

62. Liang, K.-Y and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. Biometrika, *73*, 13–22.

63. Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. Journal of the Royal Statistical Society Series B 44:226–233.

64. Mara Tableman and Jong Sung Kim, Survival Analysis Using S: Analysis of Time-to-Event Data. (2004) Chapman & Hall/CRC, Boca Raton.

65. Mc Gilchrist, C.A. (1993). REML estimation for survival models with frailty. Biometrics, 49, 221-25.

66. Miller, R.G. Survival Analysis. (1981) .Wiley & Sons, New York.

67. Minstry of finance and development, (2010), Ethiopia: 2010MGDs Report Trend and prospects for meeting MDGs by 2015,Addis Ababa Ethiopia.

68. Mutunga, C. J. (2004). Environmental determinants of child mortality in Kenya. Kenya Institute for Public Policy Research and Analysis (KIPPRA), Nairobi, Kenya.

69. Neale MC, Cardon LR . Methodology for genetic studies of twins and families. (1992) Dordrecht: Kluwer.

70. Nelson, W. (1969). Hazard plotting for incomplete failure data. Journal Quality Technology, 1, 27-52.

71. Nielsen, G.G., Gill, R.D., Andersen, P.K., Sørensen, T.I.A. (1992 . A counting process approach to maximum likelihood estimation in frailty models. Scandinavian Journal of Statistics, 19: 25–43.

72. Niragire F, A Wangombe, TNO Achia (2006). Use of the shared frailty model to identify the determinants of child mortality in Rwanda. African journals online.

73. Njagi and Purity. (2011). Urban rural differnetials of infant mortality in Kenya using a shared frailty model. Thesis , College of Humanities and Social Sciences , University of Nairobi.

74. Oakes, D. (1989). Bivariate survival models induced by frailties. Journal of the American Statistical Association, 84, 487-93.

75. Olalekan A, U., Mubashir B, U. and Ismail, Y. (2008). A Population-based study of effect of multiple birth on infant mortality in Nigeria. BMC Pregnancy and Childbirth , 41-47. http://www.biomedcentral.com/1471-2393/8/41

76. Orchard, T. and M. A. Woodbury (1972). A missing information principle: Theory and applications. In Proc. 6th Berkley Symp,Volume 1, pp.697–715.

77. Parner, E. (1998). Asymptotic theory for the correlated gamma-frailty model. The Annals of Statistics, 26(1), 183-214.

78. Parker, Schoendorf and Kiely. (2001). A comparison of recent trends in infant mortality among twins and singletons. Paediatric and Perinatal Epidemiology, 12-18. www.ncbi.nlm.nih.gov

79. Petersen, J. H. (1998). An additive frailty model for correlated lifetimes. Biometrics, 54: 646–661.

80. Peterson, J.H., Anderson, P.K. and Gill, R.D. (1996). Variance components models for survival data. Statistica Neerlandica, 50(1), 193-211.

81. Pickles, A., Crouchley, R., Simonoff, E., Eaves, L., Meyer, J., Rutter, M., Hewitt, J., Silberg, J. (1994). Survival models for developmental genetic data: age of onset of puberty and antisocial behavior in twins. Genetic Epidemiology, 11: 155–170.

82. Pietzner, D, and Wienke, A. (2010). A semiparametric estimation procedure for the bivariate correlated gamma frailty model. Institute of Medical Epidemiology, Martin-Luther-University, Germany.

83. Pison G. (1992). Twins in sub-Saharan Africa: frequency, social status and mortality. In: van der walle E, Pison G, sala-Diakanda M. mortality and society in sub-Saharan Africa. Oxford : Clarendon press , 253-278.

84. Roudsari B, Utter .G, Kernic .M, Mueller.B . (2006). Risk of early childhood injuries in twins and singletons. Journal of Early Childhood Research, vol. 4 no. 2, 121-131.

85. Salihu, H.M., Wanda, S. Aliyu, M.H., Sedjro, J.E., Bosny J. Kirby, R.S., and Alexander, G.R., Influence of nativity on neonatal survival of black twins in the United states. Ethnicity & Disease, Volume 15, Spring 2005.

86. Sastry, N. (1997). Family-Level Clustering of Childhood Mortality Risk in Northeast Brazil. Population Studies, 51, 245-261.

87. Self, S.G. and Liang, K.Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. Journal of the American Statistical Association, 82: 605-10.

88. Shi, p., Fung, W.k. (2003). Estimation of covariate effects in Cox regression models with frailties. The Indian journal of statistics, 65:56-71.

89. Steenbergen, M.R. (2006). Maximum Likelihood Programming in R. Department of Political Science, University of North Carolina, USA.

90. Stevenson.M. (2009). An Introduction to Survival Analysis. EpiCentre, IVABS, Massey University. http://epicentre.massey.ac.nz

91. Susman, A.S., (2012). Child mortality Rate in Ethiopia. Iranian J Pub Health, 3:9-19.

92. The Ethiopian Gemini Trust. Twins in Ethiopia. 2010 [cited 10 May 2010]; Available from: www.geminitrust.org/.

93. Therneau, T. M., Grambsch, P. M., and Fleming, T. R. (1990). Martingale-based residuals for survival models. Biometrika, 77: 147.160.

94. UNICEF (2010). Child mortality rate in Ethiopia falls by 40 percent. [Cited 9 October 2011]; Available from: http://www.medindia.net.

95. Vaida, F. and Xu, R. (2000). Proportional hazards model with random effects. Statistics in Medicine, 19, 3309-24.

96. Van den Berg, G.J., and Drepper, B. Inference for Shared-Frailty Survival Models with Left-Truncated Data. IZA Discussion Paper No. 6031, October 2011.

97. Vaupel, J., Manton, K., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16, 439–54.

98. Wang L. (2003). Environmental Determinants of Child Mortality: Empirical Results from the 2000 Ethiopia DHS. World Bank, Washington D.C.

99. Wienke, A. Frailty Models in Survival Analysis. (2006) Universitat Halle-Wittenberg, Medizinische Fakultat, Habilitation, 153 Seiten.

100. Wienke, A. Frailty Models in Survival Analysis. (2011) Chapman & Hall/CRC, Taylor & Francis Group.

101. Wienke, A., Lichtenstein, P. and Yashin, A.I. (2003). A bivariate frailty model with a cure fraction for modeling familial correlations in diseases. Biometrics, 59, 1178-83.

102. Wienke, A., Holm N., Christensen, K., Skythe, A., Vaupel, J.W., and Yashin, A.I. (2003). The heritability of cause-specific mortality: a correlated gamma-frailty model applied to

mortality due to respiratory diseases in Danish twins born 1870–1930. Statist. Med, 22:3873–3887.

103. Wienke, A, Christensen, K., Holm, N., and Yashin, A. I.(2002). Genetic analysis of cause of death in a mixture model of bivariate lifetime data. Statistical Modelling: 2: 1–14.

104. Wienke, A., Holm N., Skythe, A. and Yashin, A.I. (2001). The heritability of mortality due to heart diseases: A correlated frailty model applied to Dansih twins. Twin Research, 4, 266-74.

105. Wu, C. F. J. (1983). On the convergence properties of the EM algorithm. The Annals of Statistics, 1 (1): 95–103.

106. Xu, R. (2004). Proportional Hazards Mixed Models: A Review with Applications to Twin Models Metodoloski zvezki, 1(1):205-212.

107. Yashin, A.I., Vaupel, J.W. and Iachine, I.A. (1995). Correlated individual frailty: An advantages approach to survival analysis of bivariate data. Mathematical Population Studies, 5, 145-59.

108. Yashin, A.I. and Iachine, I.A. (1998). What Difference Does the Dependence Between Durations Make? Insights for Population Studies of Aging . Mathematics and Statistics Lifetime Data Analysis , 5(1):5-22.

109. Yashin, A.I. and Iachine, I.A. (1995). Genetic analysis of durations: Correlated frailty model applied to survival of Dansih twins. Genetic Epidemiology, 12, 529-38.

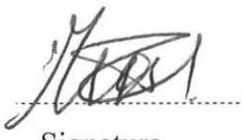
110. Yashin, A.I., Iachine, I.A., Begun, A.Z. and Vaupel, J.W. (2001). Hidden frailty: Myths and reality. Research Report 34, Department of Statistics and Demography, Odense University, Denmark.

111. Yashin, A.I. (2001). Correlated Frailty Models in the Analysis of Dependent Failure Times. Max Planck Institute for Demographic Research 18057 Rostock, Germany.

112. Yates, F. (1933). The analysis of replicated experiments when the field results are incomplete. Empire J. Exptl. Agric. 1:129–142.

DECLARATION

I, the undersigned, hereby declare that the thesis entitled "Survival analysis of under-five twins in Ethiopia: a gamma frailty modelling approach" is my original work, has not been presented for a degree in any other university and that all sources of material used for this thesis have been duly acknowledged.

Mesfin Tsegaye Hayleyesus		June 15/2013.
Student	Signature	Date

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

Dr. Emmanuel Gabreyohannes	.....	June 15, 2013
Advisor	Signature	Date