

2013•2014
FACULTY OF SCIENCES
Master of Statistics: Biostatistics

Masterproef
Survival modelling of rehospitalization in telemonitored chronic heart failure patients

Promotor :
dr. Edmund NJAGI

Supervisor :
Prof. dr. Geert MOLENBERGHS
Prof. dr. Paul DENDALE

Transnational University Limburg is a unique collaboration of two universities in two countries:
the University of Hasselt and Maastricht University.



Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek

Richard Njilefac
*Master Thesis nominated to obtain the degree of Master of Statistics , specialization
Biostatistics*



2013•2014
FACULTY OF SCIENCES
Master of Statistics: Biostatistics

Masterproef

Survival modelling of rehospitalization in
telemonitored chronic heart failure patients

Promotor :
dr. Edmund NJAGI

Supervisor :
Prof. dr. Geert MOLENBERGHS
Prof. dr. Paul DENDALE

Richard Njilefac

*Master Thesis nominated to obtain the degree of Master of Statistics , specialization
Biostatistics*

Acknowledgements

I will begin by thanking the almighty God who gave me life and has been the source of my strength from day one of this master program until the completion of this thesis.

My sincere appreciation goes to my supervisor, Dr. Edmund Njeru Njagi, who, from the beginning to the end, stimulated my thinking, provided substantive suggestions and was always available at my calls. Also to Prof. Dr. Geert Molenberghs and Prof. Dr. Paul Dendale who read and made valuable corrections, I say a big thank you.

I am highly indebted to the family that brought me to life and moulded me up, instilled in me the value and desire for education and discipline and have always provided financial and moral support to see me succeed. In this regard, I say a big thank you to my late dad, Johnson Nkemaka, my mom, Magdalene Emenkeng, my brothers and sisters and the entire family. To my second parents, Commissaire Tatu Nico Atem (Chief Formin) and wife Tatu Patricia, I can't cease to acknowledge and appreciate the orientation, material and moral backup you have always provided me.

I acknowledge and appreciate the entire Censtat staff at Hasselt University for their commitment in disseminating statistical knowledge, building our capacity in handling multiple tasks and yet meeting deadlines. You have indeed made me a proud statistician.

I also wish to say thank you to all my friends and colleagues at Hasselt University especially Blaise Ngantcha, Silenou Bernard, Kouam Nouria, Ngu Blessing, Alemgu Nico, Efuetao Charles, Pascal Atemnkeng, Karin Daniels and Marcellus Atanga just to name a few, with whom we repeatedly shared study projects and extracurricular activities. Our discussions and your advice made a significant contribution to my success throughout these years.

Abstract

Chronic heart failure is a common and costly medical condition that is characterised by repeated patient hospitalisation. This condition arises when the heart is unable to pump enough blood to supply the needs of the body. The cost burden of this illness has been estimated at billions of dollars with the major contributing factor being the need for rehospitalisation. As a strategy in heart failure management, remote patient monitoring, known as telemonitoring, enables clinicians to predict rehospitalisation so that in-time intervention decisions can be made. Aiming at detecting baseline patient characteristics associated with rehospitalisation in these patients, the so-called one parameter shared gamma frailty model was applied, assuming an exponential and Weibull distributed baseline hazards. In this way, the clustering or correlation resulting from the repeated events (rehospitalisation) within patients was taken into account. The time ordinality observed in repeated events was accounted for via two formats: the gap time and the calendar time representation. The patient's heart rhythm was found to be a significant predictor (p -value=0.0231) of rehospitalisation at 5% level of significance. However, these results have a limitation due to the small sample size used in the study.

Keywords: Chronic heart failure, frailty model, calendar time representation, gap time representation.

Contents

Acknowledgements.....	i
Abstract.....	ii
List of abbreviations.....	iv
List of figures.....	iv
1. Introduction.....	1
1.1 The Chronic Heart Failure Data.....	3
2. Statistical Methodology.....	4
2.1. Exploratory Data Analysis.....	4
2.2. Statistical Analysis.....	4
2.2.1. Survival Analysis.....	4
2.2.2. Proportional Hazards Model.....	5
2.2.3 Multivariate (or Clustered) Survival Data Analysis.....	6
2.2.4. Model Selection.....	13
2.2.5. Statistical Software.....	13
3. Results.....	14
3.1. Exploratory Data Analysis.....	14
3.2. Statistical Analysis.....	15
6. Discussion.....	18
7. Conclusion.....	19
8. Limitations and Recommendations.....	19
9. References.....	20
10. Appendix.....	25
10.1. Appendix A.....	25
10.2 Appendix B.....	30

List of abbreviations

CHF	: Chronic Heart Failure
HF	: Heart Failure
LVEF	: Left Ventricular Ejection Fraction.
NT-ProBNP	: N-terminal pro brain natriuretic peptide.
NYHA	: New York Heart Association
PH	: Proportional Hazards
S.E	: Standard Error.
TM	: Telemonitoring

List of figures

Figure 3.1	: Kaplan-Meier survival estimate for time to first rehospitalisation.....	15
------------	---	----

List of tables

Table 1.1	: Description of variables used in the CHF data set.....	3
Table 3.2.1	: Results of the frailty model with all main effects and Weibull baseline hazard.....	16
Table 3.2.2	: Results of the frailty model with all main effects and exponential baseline hazard.....	16

1. Introduction

Heart failure (HF) describes a medical condition in which the heart is unable to provide sufficient pumping action to maintain blood flow to meet the needs of the body. It is a common, serious and costly health problem which is usually characterised by recurrent hospitalisations. Initially thought to be due to fluid overload and/or worsening of renal function (Dendale *et al.*, 2011) (called the “cardiorenal model of heart failure (Douglas *et al.*, 2005), a later realisation has also associated heart failure to excessive peripheral vasoconstriction and reduced cardiac output (known as the “cardiocirculatory” or “hemodynamic” model of heart failure) (Packer, 1992; Douglas *et al.*, 2005). Heart failure is initiated by what is known as an “index event” which leads to loss of functioning of cardiac myocytes as a result of heart muscle damage. The ability of the myocardium to generate force is lost and as such the heart is unable to contract normally during its pumping action (Packer, 1992).

According to estimates from the European Heart Failure Association, 26 million people have HF worldwide and 3.6 million people are newly diagnosed with HF every year in Europe alone (López, 2011).

Gheorghide *et al.* (2013) estimated that one million hospitalisations due to heart failure occur annually in the United States. Even though much progress is being made in reducing mortality in these patients, rates of rehospitalisation have kept on rising, approaching 30% within 60 to 90 days of discharge.

The economic cost of HF has been estimated at billions of dollars per year, with the need for repeated hospitalisation being the highest contributing factor that directs cost associated with the disease (López, 2011). Therefore, investigating prognostic factors and proposing solutions that can remedy these rehospitalisations in CHF patients is of great importance to health care providers. This will go a long way to reduce cost burden on both the government and the individual patients.

In the heart failure literature, many risk factors have been outlined to be associated with rehospitalisation in HF patients. It is important to keep in mind that there exists an inexhaustible list of such factors spanning from demographic to clinical and environmental factors. Therefore, aiming at investigating all these factors may be an unachievable venture. In this vein, among other authors, Anderson and Steinberg (1985), and Thomas (1996)

reported gender and age as potential demographic predictors of rehospitalisation in HF patients, with older patients and females more likely to be rehospitalised due to HF disease. Many studies have also considered the left ventricular ejection fraction (LVEF), which is a measure of how much blood is pumped out of the left ventricle of the heart (the main pumping chamber). It has remained an inconsistent predictor, with some studies reporting its insignificance (Cheng *et al.*, 2001; Logeart *et al.*, 2004; Goonewardena *et al.*, 2008), while others suggest lower LVEF patients are more likely to be readmitted (Dokainish *et al.*, 2005; Gackowski *et al.*, 2004). A heart disease severity factor such as New York Heart Association class of the disease has also been reported to be a prominent risk factor in predicting rehospitalisation in this type of patients. It depicts a patient's functional capacity and does classify patients' HF according to the severity of their symptoms. In this regard, patients are placed in one of four ordinal categories depending on how much they are limited during physical activity. Smith *et al.* (2000) and Yamokoski *et al.* (2007) reported higher NYHA class patients as more likely to be rehospitalised. Heart rhythm (otherwise known as sinus rhythm) is a measure of the normal heart beat. Its effect on rehospitalisation of HF patients has been studied. Formiga *et al.* (2004) and Murakami *et al.* (2012) found no association between risk of rehospitalisation and heart rhythm meanwhile Korte *et al.* (2000) and Linssen *et al.* (2011) reported a higher risk of rehospitalisation in patients with abnormal heart rhythm (precisely the abnormality known as atrial fibrillation).

Clinical parameters like natriuretic peptides (for example N-terminal-pro-brain natriuretic peptide (NT-proBNP)) are other important prognostics factors that have found significant grounds in the prognosis of heart failure in the heart failure literature. NT-proBNP is an inactive 76 amino acid peptide that is usually secreted along with a brain natriuretic peptide (BNP) by the heart ventricles. Their secretion is usually in response to the excessive stretching activity of the heart muscle cells (cardiomyocytes). It serves as a good target for diagnostic blood testing due to its long biological half-life. Therefore, in circumstances of excessive stretch of cardiomyocytes as the case maybe under certain heart failures, one would expect higher levels of these peptides in blood. Bettencourt *et al.* (2004) and Mayur (2009) showed NT-ProBNP to be associated with the risk of rehospitalisation in HF patients with high NT-proBNP level patients having more risk of rehospitalisation compared to lower NT-proBNP level patients.

The main objective of this study was to investigate the relationship between baseline patient characteristics and the risk of rehospitalisation in telemonitored chronic heart failure patients.

Section 1.1 continues with a description of the data set. In Section 2, a general discussion of survival data analysis methodology is presented, and, in Section 3, results obtained after implementing these methods are shown. Section 4 and 5 continue with a discussion of the results and conclusive remarks respectively, while in Section 6, the limitations of this study, and recommendations, are discussed.

1.1 The Chronic Heart Failure Data

These data came from a study conducted in Belgium between 2008 and 2010 (Dendale *et al.*, 2011), where chronic heart failure (CHF) patients were followed for a period of 6 months. The goal of the study was to investigate whether observing patients by telemonitoring would reduce mortality and rate of rehospitalisation. Telemonitoring (TM) involved providing a set of electronic appliances to the patients upon hospital discharge to collect measurements of heart rate, blood pressure and weight on a daily basis. These measurements were then remotely transferred to the heart failure clinic and general practitioners for assessment and possible intervention. There was a total of 160 CHF patients, 80 of whom were followed by means of TM, while 80 received the usual care.

In this study, our interest is to investigate the relationship between the baseline patient characteristics and the risk of rehospitalisation for the telemonitored group. As such, analyses will be conducted on data from the 80 telemonitored patients. These data have also been analysed by Njagi *et al.* (2013a, b) in different contexts. The variables used in the data set are as described in Table 1.1.

Table 1.1 Description of variables used in the CHF data set

Variable	Type	Description
Ptid	Character	Patient identification number.
Status	Categorical	Censoring indicator (0=censored, 1=event).
Day (response)	Quantitative	Day at which an event or censoring occurred.
Sex	Categorical	Gender of the patient (0=female, 1= male).
NYHA	Quantitative	New York Heart Association class.
LVEF	Quantitative	Left ventricular ejection fraction.
Age	Quantitative	Age of the patient at entry into study.
Heartrym	Categorical	Heart Rhythm (0=normal, 1=not normal).
NTproBNP	Quantitative	N-Terminal-pro-Brain Natriuretic Peptide.

2. Statistical Methodology

2.1. Exploratory Data Analysis

To gain some insight into the data structure and patterns within the data, an exploratory data analysis was conducted. Summary statistics were obtained for the various variables in the data set, and, to understand the shape of the survival function, the Kaplan-Meier curve was used.

2.2. Statistical Analysis

2.2.1. Survival Analysis

Survival analysis (as commonly referred to in the biological/health sciences) is a branch of statistics that defines a set of methods in analysing data where the outcome variable is time to occurrence of some event of interest. These methods are widely applied in diverse fields such as engineering sciences where it is often referred to as reliability analysis (or failure time analysis or in economics as duration analysis, etc. In biological/health sciences, examples of events could be death, occurrence of a disease, hospitalisation, relapse of a given disease, or cure from a given disease, while in engineering sciences, events could include failure of a mechanical system, where as in economic studies, it could be time to find a new job.

A feature central to survival analysis differing from classical statistics is censoring. This is a type of missingness where some observations are not complete or are only partially observed. Hence survival methods are constructed such that these partial observations are taken into account during analysis. In the light of censoring, we can differentiate three types of censoring: right censoring, left censoring and interval censoring. Right censoring is the more common form of censoring and it occurs when only the lower bound of the time of interest is known. That is, we know that a subject has survived up to a particular time and thereafter, the subject is no longer followed up. In left censoring, the subjects have already experienced the event of interest at the time they are entered into the study. In interval censoring, the event of interest is only known to occur within a certain time interval, but the exact time is not known.

A second feature at times present in some survival studies is that of truncation. Two forms of truncation can be distinguished; left truncation which occurs when subjects enter a study at for example, a particular age (a precondition) and are followed from this delayed entry time till occurrence of an event or censoring, and right truncation which occurs when only subjects who have experienced the event of interest are included in the study. It is important to

remember that there are many types of censoring schemes within left and right censoring and each type will lead to a different survival likelihood function. However, Klein and Moeschberger (2003) have highlighted that even though the likelihood function is unique for each type of censoring, there exist a common approach to be used in constructing it.

2.2.2. Proportional Hazards Model

The proportional hazards (PH) model is the most popular model used in survival data analysis (Duchateau and Janssen, 2008). Proposed by Cox (1972), the model consist of two parts: the underling hazard function (also known as the baseline hazard function), which describes how the risk of an event changes over time at baseline levels of the covariate, and the effect parameters, which describe how the hazard function varies in response to risk factors (or covariates). Borrowing the notation as described by Duchateau and Janssen (2008), the PH model can simply be represented as:

$$h_T(t) = \psi h_C(t) , \quad [1]$$

where the constant, $\psi = \frac{h_T(t)}{h_C(t)}$ defines the hazard ratio when we consider a simple single covariate example of a treated versus untreated (control) group of patients. $h_T(t)$, is the hazard function of patients in the treatment group and $h_C(t)$ is the hazard function of patients in the control group. This relationship therefore implies that the ratio of hazards of treated versus untreated subjects does not change over time. This is known as the proportional hazard assumption.

Because the hazard ratio must be positive, a common model for ψ is $\exp(\beta)$. When risk factors are to be included in the analysis, the PH model is then written as:

$$h_i(t) = h_0(t) \exp(\mathbf{x}_i' \boldsymbol{\beta}) , \quad [2]$$

where $h_0(t)$ is the baseline hazard function that corresponds to the hazard function of a subject whose covariate information \mathbf{x}_i is $\mathbf{0}$; $h_i(t)$ is the hazard function of the i^{th} subject and $\boldsymbol{\beta}$ is the vector of parameter estimates for the predictors. When the baseline hazard function is left unspecified, the model is termed a semi-parametric PH model, whereas when it is specified to follow a given distribution, it is said to be a parametric PH model.

2.2.3 Multivariate (or Clustered) Survival Data Analysis

A higher level of complexity in survival analysis arises when data is clustered. The observational units (subjects) in which event times are observed may be clustered, or event times may be clustered on the observational units (otherwise known as repeated events as in this study). Several analysis techniques have been proposed to handle such data; they include the fixed effects Cox model, stratified Cox model, copula model, the marginal model, and the frailty model.

In the fixed effect model, the cluster effect is modelled via a fixed effect parameter as can be described in the following model:

$$h_{ij}(t) = h_o(t) \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta} + c_i), \quad [3]$$

where c_i is the fixed effect for the i^{th} cluster, $h_{ij}(t)$ is the hazard function for observation j of cluster i , $h_o(t)$ is the baseline hazard common to all subjects in the population, \mathbf{x}_{ij} is the vector of covariates and $\boldsymbol{\beta}$ is the vector of parameter estimates for fixed effects (Duchateau and Janssen, 2008). As stated by Glidden and Vittinghoff (2004), when the number of clusters is small relative to the sample size, model [3] is an attractive approach to account for the clustering, especially if the cluster effects are of intrinsic interest. However, large numbers of clusters relative to the sample size can be problematic. The asymptotics break down when the number of clusters tend to infinity as observations increase and consequently, parameter estimation is quite inefficient and model interpretation is less natural (Duchateau and Janssen, 2008). An alternative approach is to use the semi-parametric stratified Cox model whereby a separate unspecified baseline hazard is attributed to each cluster. This can be represented as:

$$h_{ij}(t) = h_{i_o}(t) \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}), \quad [4]$$

with $h_{i_o}(t)$ the baseline hazard for cluster i . This model has a flexibility advantage over the fixed effect model [3] in that the baseline hazard can evolve independently over time within each cluster (Duchateau and Janssen, 2008). Therefore, the hazard functions could have different shapes with some or all unequal (Glidden and Vittinghoff, 2004). In contrast, in the fixed effect model [3], the baseline hazard is restricted to be of the form $h_o(t) \exp(c_i)$, with c_i the constant specific effect for cluster i . On the other hand, the stratified Cox model has some setbacks that limit its application. As discussed by Duchateau and Janssen (2008), a cluster will only contribute to the partial likelihood expression of the model if an event for a subject

is observed while the other subject is still at risk. In a repeated event setting, this therefore implies that a subject will only contribute to the partial likelihood if it has at least two events. Glidden and Vittinghoff (2004) further expatiated that, there is considerable loss of information in the application of this model as in the unstratified Cox model since intercluster comparisons are impossible. Hence, Duchateau and Janssen (2008) clearly stated that in many practical situations this model will be inefficient.

A third approach to account for the cluster effect in survival data is to use the copula model. This model is often considered to be a two stage model. In the first stage, a marginal survival function for each observation in a cluster (or for each event in a subject in a repeated event setting) is obtained ignoring the clustering. In the second stage, a copular function is then used to link the population survival functions to form a joint survival function (Fress *et al.*, 1996) which is expressed as:

$$S_{\mathbf{x},p}(t_n) = S_{\mathbf{x},p}(t_1, \dots, t_n) = C_{\boldsymbol{\theta}}(S_{\mathbf{x}_1,p}(t_1), \dots, S_{\mathbf{x}_n,p}(t_n)), \quad [5]$$

where $\mathbf{x} = (x_1^t, \dots, x_n^t)^t$ is the covariate information, n is the number of observations, $S_{\mathbf{x},p}(t)$ is the join population survival function and $C_{\boldsymbol{\theta}}$ is the copula function with parameter vector $\boldsymbol{\theta}$.

The marginal model approach describes another handy method in the analysis of clustered survival data. In this method, clustering is completely not accounted for in the first step. The modelling proceeds as if the event times are independent on each other even when they belong to the same cluster (or subject in a repeated event setting). This approach consists of two stages. In the first stage, an independence working model (IWM) [6] is fitted to the data.

$$h_{ij}(t) = h_o(t) \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}), \quad [6]$$

where $h_{ij}(t)$, $h_o(t)$, \mathbf{x}_{ij} and $\boldsymbol{\beta}$ are described as in model [3]. Duchateau and Janssen (2008) also provided a clear demonstration to show that the parameter estimates from the IWM are consistent even when clustering is not accounted for. However, the inverse of the information matrix of $\hat{\boldsymbol{\eta}}$ (a vector of parameter estimates) is not a consistent estimator of the asymptotic variance-covariance matrix when the correlation of survival times is ignored as in the IWM (Duchateau and Janssen, 2008). At the second stage, a sandwich estimator (for example the jackknife estimator) is used to estimate the variance of $\hat{\boldsymbol{\eta}}$.

A more widely used method in modelling survival data with clustering is the so call shared frailty model (McGilchrist and Aisbett, 1991). This is an alternative formulation of the random effects model in classical clustered data. A “frailty” is an unobservable random effect shared by subjects within a cluster (or by events within a subject in a repeated event setting), (Klein and Moeschberger, 2003). Clayton (1978) first applied this model in a multivariate situation in his study of chronic disease incidence in families. In this model, the common random effect acts multiplicatively on the hazard rates of all subjects within a cluster (Klein and Moeschberger, 2003). The model can be seen as an extension of the Cox proportional hazards regression model (Cox, 1972) where we assume that the hazard rate for the j th subject in the i th cluster given the frailty, is of the form:

$$h_{ij}(t) = h_o(t) \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta} + w_i), \quad [7]$$

where $h_{ij}(t)$ is the conditional hazard function for the j^{th} subject of cluster i conditional on w_i , $h_o(t)$ is the baseline hazard usually treated parametrically, \mathbf{X}_{ij} is the vector of covariates, $\boldsymbol{\beta}$ is the fixed effects vector of dimension p and w_i is the random effect for cluster i . This model can as well be rewritten as

$$h_{ij}(t) = h_o(t) \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}), \quad [8]$$

where $\varphi_i = \exp(w_i)$ is called the frailty for the i^{th} cluster and the φ_i 's, $i=1, \dots, s$, are the actual values of a sample from some density f_φ . The parsimony that comes along with this model has made it very popular for use in clustered failure time data (Sujit and Dipak, 1998).

Usually, the choice for a parametric baseline hazard lies in exploiting the classical maximum likelihood techniques in parameter estimation, since the marginal survival likelihood will be fully parametric (Duchateau and Janssen, 2008). The Weibull and the Exponential distribution was used for this purpose. While the exponential baseline hazard function simply assumes the hazard rate to be constant over time, the Weibull is more flexible. The Weibull hazard function can be monotone increasing, decreasing, or constant (Klein and Moeschberger, 2003). Furthermore, according to Duchateau and Janssen (2008), assuming the Weibull distribution for the baseline hazard function is usually an appropriate assumption in most practical situations. Nonetheless, there exist other distributions like the Gompertz, the Loglogistic and the lognormal distributions that have also been suggested. But in the context of parametric frailty models where the marginal survival function is analytically derived, the

major limiting factor behind the use of these distributions for baseline hazard functions lies in the mathematical complexity of the expressions.

As discussed earlier in this section, in parametric shared frailty models the distribution of random affects (the frailty) is required to be specified. The one parameter gamma distribution is the most commonly assumed distribution in the clustered survival data literature. The motivation for the choice of this distribution as discussed by Duchateau and Janssen (2008) lies essentially in the fact that it is easy to analytically integrate out the frailty term in the conditional survival likelihood, resulting in an explicit and simple expression for the marginal survivor function. This integration is however still possible with other distributions suggested in the literature such as: the inverse Gaussian distribution, the compound poisson distribution, etc. (Duchateau and Janssen, 2008; Hanagal and Dabade, 2013) they lead to much more complex expressions.

An important feature of recurrent events, such as the recurrent rehospitalisation in our data, is the time ordering of events. If a patient had two rehospitalisations, then the second can only occur sometime after the first has occurred. There therefore exists natural ordering in event times. Moreover, it is very possible that previous hospitalisations could have an influence in the subsequent rehospitalisation times. To account for this feature in frailty models, two forms of time scales are used (Duchateau and Janssen, 2008).

2.2.3.1 Gap Time Representation

In the gap time representation, the time at risk for an event is always reset to start at zero after every event. However, the length of time at risk corresponds to the time since the end of the previous event (or entry to the study in the case of the first event) to the time of the particular event. The frailty model [8] becomes:

$$h_{ij}(t) = h_o(t - y_{ij1})\varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) \quad \text{for} \quad y_{ij1} \leq t \leq y_{ij2}, j = 1, \dots, n_i$$

And the cumulative hazard $H_{ij}(y_{ij})$ for the risk period j of the i^{th} subject is given by

$$H_{ij}(y_{ij}) = \int_0^{y_{ij2} - y_{ij1}} h_{ij}(t) dt,$$

where $\mathbf{y}_{ij} = (y_{ij1}, y_{ij2})^t$ refers to the start and the end of the risk interval j .

The conditional density for subject i , given frailty term, is then given by:

$$f(t/\varphi_i) = h_{ij}(t/\varphi_i) s(t),$$

where $s(t)$ is the survival function.

For a Weibull baseline hazard, the expression becomes:

$$f(t/\varphi_i) = [\lambda \rho (t - y_{ijt})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_0^z \lambda \rho (t)^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right), \quad [9]$$

where δ_{ij} is the censoring indicator and $z = y_{ij2} - y_{ij1}$ is the time interval for an event from the start of at risk period j of subject i to when the event occurred.

For an exponential baseline, the expression changes to:

$$f(t/\varphi_i) = [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_0^z \lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \quad [10]$$

To obtain the marginal density functions, assuming a one parameter gamma density for the frailty distribution, we integrate the product of the conditional densities [9] and [10] and the gamma density for the frailty distribution as follows:

$$f(t) = \int_0^\infty [\lambda \rho (t - y_{ijt})^{\rho-1} \varphi_i \exp(\boldsymbol{\beta} \mathbf{x})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\boldsymbol{\beta} \mathbf{x}) z^\rho) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i; \quad [11]$$

for the Weibull case, and:

$$f(t) = \int_0^\infty [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \quad [12]$$

for the exponential baseline hazard.

The distribution $f_{\varphi_i}(\varphi) = \frac{\varphi^{1/\theta-1} \exp(-\varphi / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)}$ is the one parameter gamma density with mean 1 and variance θ .

Integration and simplifications result in the expressions:

$$f(t) = \frac{[\lambda \rho (t - y_{ijt})^{\rho-1} \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho + 1/\theta)^{(1/\theta + \delta_{ij})}}, \quad [13]$$

when the baseline hazard is Weibull distributed, and:

$$f(t) = \frac{[\lambda \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z + 1/\theta)^{(1/\theta + \delta_{ij})}}, \quad [14]$$

when the baseline hazard is exponential, and Γ is the gamma function.

Expressions [13] and [14] are marginal density functions for one parameter gamma frailty models assuming the Weibull and exponential baseline hazards respectively.

Details on integration and simplifications to obtain these marginal density functions are provided in Appendix A.

2.2.3.2. Calendar Time Representation

In this form of time representation, the length of the at risk period for an event is the same as in the gap time representation; that is, the time since the end of the previous event (or entry into the study in the case of the first event), to the time when the particular event occurred. However, the start of the at risk period is not reset to zero as was the case under gap time representation, but is rather the actual time since entry into the study. In this case, the frailty model [8] is now described as:

$$h_{ij}(t) = h_o(t) \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) \quad \text{for} \quad y_{ij1} \leq t \leq y_{ij2}, j = 1, \dots, n_i,$$

where t in this case refers to the time since entry into the study. The cumulative hazard $H_{ij}(y_{ij})$ for at risk period j of the i^{th} subject is now given by

$$H_{ij}(y_{ij}) = \int_{y_{ij1}}^{y_{ij2}} h_{ij}(t) dt,$$

with $\mathbf{y}_{ij} = (y_{ij1}, y_{ij2})^t$ as described in section 2.2.3.1.

The conditional density for subject i , given the frailty term, is then given by:

$$f(t/\varphi_i) = h_{ij}(t/\varphi_i) s(t)$$

where $s(t)$ is the survival function.

For a Weibull baseline hazard, the expression becomes:

$$f(t/\varphi_i) = [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_{y_{ij1}}^{y_{ij2}} \lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \quad [15]$$

where δ_{ij} is the censoring indicator, y_{ij1} and y_{ij2} are the start and end of the at risk period j for an event in subject i respectively.

For an exponential baseline hazard, expression [10] becomes:

$$f(t/\varphi_i) = [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_{y_{ij1}}^{y_{ij2}} \lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \quad [16]$$

To obtain the marginal density function, assuming a one parameter gamma density for the frailty distribution, we integrate as follows:

$$f(t) = \int_0^{\infty} [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) (y_{ij2}^{\rho} - y_{ij1}^{\rho})) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \quad [17]$$

when the baseline hazard is Weibull, and

$$f(t) = \int_0^{\infty} [\lambda \varphi_i \exp(\boldsymbol{\beta} \mathbf{x})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\boldsymbol{\beta} \mathbf{x}) (y_{ij2} - y_{ij1})) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \quad [18]$$

when the baseline hazard is exponential. Integrating and simplifying expressions [17] and [18] results in the expressions:

$$f(t) = \frac{[\lambda \rho t^{\rho-1} \exp(\boldsymbol{\beta} \mathbf{x})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda \exp(\boldsymbol{\beta} \mathbf{x}) (y_{ij2}^{\rho} - y_{ij1}^{\rho}) + 1/\theta)^{(1/\theta + \delta_{ij})}} \quad [19]$$

for the Weibull, and

$$f(t) = \frac{[\lambda \exp(\boldsymbol{\beta} \mathbf{x})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda \exp(\boldsymbol{\beta} \mathbf{x}) (y_{ij2} - y_{ij1}) + 1/\theta)^{(1/\theta + \delta_{ij})}}, \quad [20]$$

for the exponential case, where Γ is once again the gamma function.

Again, details on the integration and simplifications of expressions [19] and [20] are provided in Appendix A.

An interesting observation to note is that when we replace $z = y_{ij2} - y_{ij1}$ in expression 14, expressions 14 and 20 are exactly the same. This implies that, for the exponential baseline hazard, the calendar and gap time representations are not different. This is related to the fact that the exponential distribution is memoryless (Duchateau *et al.*, 2003).

For the entire analysis, the common assumption of right censoring and that the censoring was non-informative (that is, independent of event times) was adopted.

2.2.4. Model Selection

In a model building process in survival analysis, many approaches have been suggested. These include the standard automatic search procedures like the backward, forward and stepwise procedures.

In this study, the backward selection procedure was applied. This selection procedure proceeded by fitting a model with all the main effects and, at each step, the predictor with the most insignificant p-value was dropped, until no predictor could be dropped from the model, assuming a 5% level of significance.

2.2.5. Statistical Software

Statistical analysis was conducted in SAS version 9.3 while R version 3.0.2 was used to obtain Kaplan-Meier plot. Specifically, in SAS 9.3, the NLMIXED procedure was used to fit the frailty models for both calendar and gap time representations. Of notable attention in the NLMIXED procedure is the use of starting values and various optimisation technics that have serious impact in achieving model convergence and even inference (Kiernan *et al.*, 2012). In this regard, starting values for the various parameters in each frailty model were obtained by fitting a similar model in PROC LIFEREG and extracting the parameter estimates. The SAS LIFEREG procedure fits parametric univariate survival models (that is under independence assumption of event times) using the loglinear model representation. Since the frailty models were of proportional hazards nature, transformation of the parameters to the proportional hazards representation was done as discussed by Duchateau and Janssen (2008) before use. For homogeneity purposes, the Newton-Raphson optimisation technique was used throughout the analysis. Furthermore, to ease model convergence, all quantitative covariates and the response variables were rescaled to values between zero and one; this was done by dividing with the largest observation of the respective variable. All statistical tests were conducted at 5% level of significance.

3. Results

3.1. Exploratory Data Analysis

There were 80 telemonitored CHF patients, of which 50 were males, and 30 were females. The average age of the patients was 76(9.67) years, with a minimum and maximum of 46 and 95 years, respectively. The average patient fitness measure (NYHA) was 3.06(0.49) with a minimum of 2 and maximum of 4. The mean left ventricular ejection fraction (LVEF) was 35.57(15.45) %, with a minimum 12.5% and a maximum of 80%. There were 44 patients with a normal heart rhythm, and 36 with abnormal heart rhythm. For the cardiac muscle fibre stretch measure, the average N-Terminal-pro-Brain Natriuretic Peptide (NT-proBNP) was 4993.22(6835.61), with a minimum of 16 and 37690.00 as the maximum. Four patients had missing measurements for the NT-proBNP variable; these patients were not included in the analyses.

Throughout the study, 64 patients did not get any rehospitalisation (censored patients) and 16 patients were rehospitalised. Of these 16, one had 3 rehospitalisations, two had 2 rehospitalisations, and the remaining 13 had one rehospitalisation each. These made a total of 20 rehospitalisations (events) in the entire study period.

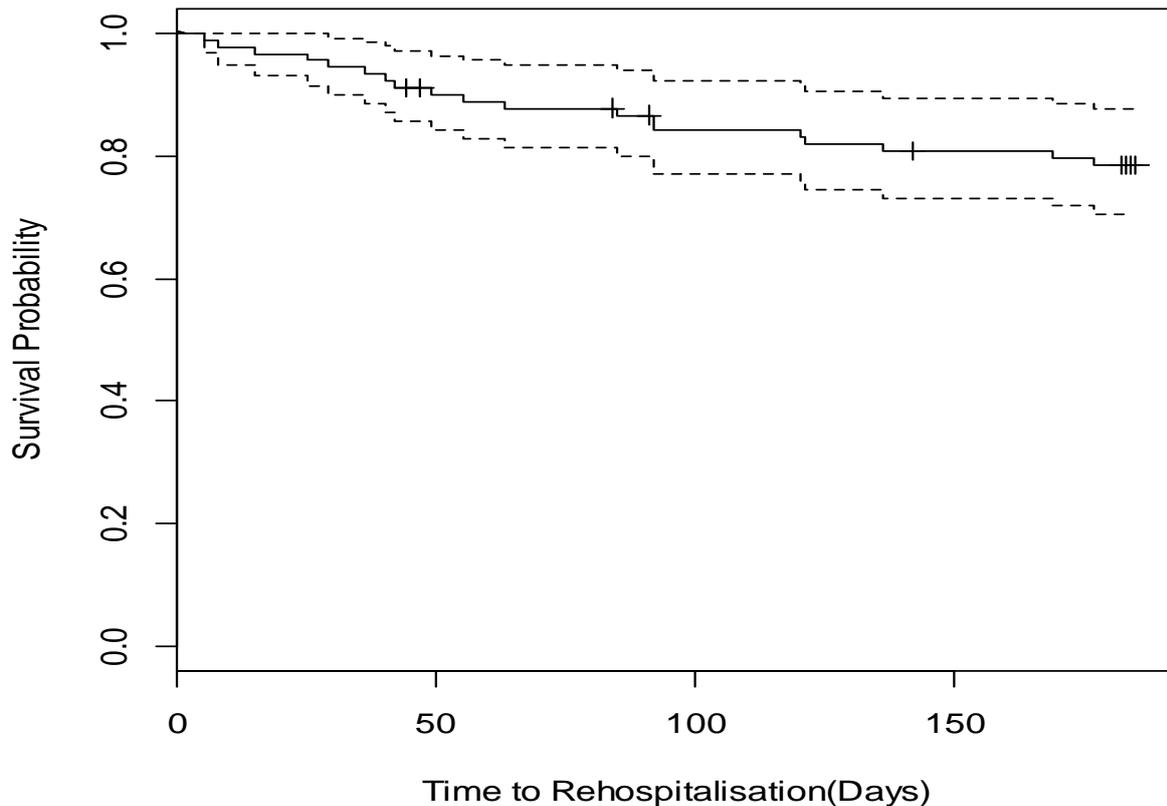


Figure 3.1 Kaplan-Meier survival estimate for time to rehospitalisation

Figure 3.1 presents the Kaplan-Meier survival estimate with 95% confidence interval for time to rehospitalisation. The curve has jumps at event times. The probability of patients staying without rehospitalisation post discharge slowly decreased with time (days post discharge). However, it is important to note that this plot is non-parametric and also does not take the repeated nature of events in subjects into consideration. Therefore, one should avoid over-interpreting it.

3.2. Statistical Analysis

Tables 3.2.1 and 3.2.2 show the results obtained from fitting the frailty model with all main effects assuming a Weibull and exponential baseline hazard respectively, considering both calendar and gap time representations. All first order interactions could not be investigated because of model non-convergence.

Table 3.2.1 Results of the frailty model with all the main effects and Weibull baseline hazard

Parameter	Calendar time Representation			Gap time representation		
	Estimate	Standard Error	p-value	Estimate	Standard Error	p-value
Age	-7.4000	6.3107	0.2440	-7.4474	6.7873	0.2754
Sex	-1.9724	1.5039	0.1930	-1.5324	1.3058	0.2437
NYHA	-1.9873	6.3298	0.7543	-4.3436	6.9613	0.5342
LVEF	-5.1255	6.4497	0.4289	-1.7314	5.0745	0.7337
NTproBNP	7.2095	5.1241	0.1628	8.9236	5.8686	0.1318
Heart Rythm	4.1389	1.7904	0.0231	4.3180	2.5011	0.0877
λ	6764.8900	23886	0.7777	4434.55	26040	0.8652
ρ	1.9655	0.6937	0.0057	1.6351	0.6940	0.0206
θ	15.0981	8.1367		13.7234	9.6440	
AIC	89.2000			91.1000		
-2logLik	71.2000			73.1000		

-2logLik= -2 Log Likelihood , AIC= Akaike's Information Criterion

Table 3.2.2 Results of the frailty model with all main effects and exponential baseline hazard

Parameter	Calendar time representation			Gap time representation		
	Estimate	Standard Error	P-Value	Estimate	Standard Error	P-Value
Age	-4.3033	5.6290	0.4466	-4.3033	5.6290	0.4466
Sex	-1.0584	0.9452	0.2657	-1.0584	0.9452	0.2657
NYHA	-1.2833	4.1285	0.7566	-1.2833	4.1285	0.7566
LVEF	-1.1975	2.9812	0.6889	-1.1975	2.9812	0.6889
NTproBNP	5.7936	3.6490	0.1158	5.7936	3.6490	0.1158
Heart Rythm	2.3325	1.2219	0.0594	2.3325	1.2219	0.0594
λ	20.9667	112.7400	0.8529	20.9667	112.7400	0.8529
θ	6.3364	2.8589		6.3364	2.8589	
AIC	90.4000			90.4000		
-2logLik	74.4000			74.4000		

-2logLik= -2 Log Likelihood , AIC= Akaike's Information Criterion

From Table 3.2.1, a goodness of fit comparison of the calendar and gap time frailty models with the assumed Weibull baseline hazard can be done. This comparison is not possible using the likelihood ratio test since the models are not nested (Duchateau and Janssen, 2008). Therefore, the comparison was done using the Akaike's information criterion (AIC); this gave a lower value for the calendar time model, providing evidence that this model had a better fit to the data. A similar comparison from Table 3.2.2 for the frailty models with the assumed exponential baseline hazard function showed that the two model fits were equal. The entire results of the two models were exactly the same. This was in line with the observation stated in Section 2.2.3.2, where similar expressions for the marginal density functions were noted for both the calendar and the gap time frailty models, under the exponential baseline. A further goodness of fit comparison of the frailty models between the two baseline hazards (Weibull and exponential) was now possible using the likelihood ratio test, since the exponential model is nested in the Weibull model. The exponential distribution is a special case of the Weibull distribution; the former results when the shape parameter (ρ) in the latter is constrained to one. Therefore, the likelihood ratio test statistic is chi-square distributed with degree of freedom one. The test ($H_0: \rho=1$) revealed a chi-square test statistic of $74.4000-71.2000=3.2000$, with $P(\chi^2 \geq 3.2) = 0.0736$, which was borderline significant. This provided evidence that the frailty model with the Weibull baseline hazard under the calendar time format provided better fit than the exponential counterpart. Attempts for model reduction to achieve a more parsimonious model by using the backward model selection as described in Section 2.2.4 could not lead further model reduction. Therefore, the conclusion was that heart rhythm was the lone significant covariate with a relative risk (or hazard ratio) of $\exp(4.1389) = 62.7338$ with a 95% confidence interval of $\exp(4.1389 \pm 1.96 * 1.7904) = [1.8770, 2096.6250]$. This implies that the risk of rehospitalisation is about 63 times higher in a patient with an abnormal heart rhythm compared to when that same patient has a normal heart rhythm given that all other covariates were held constant. The estimate for ρ is equal to 1.9655 which is substantially larger than one. The hazard function increases with time. This means that the rehospitalisation event rate increases with time since entry into the study. However, it is clear that the confidence interval for the relative risk is very wide. This implies a high degree of uncertainty in the estimation. In the same vein, standard errors of the parameter estimates for all models and in both time formats are also large, implying low precision in parameter estimations. Further still, the estimate for the scale parameter (λ) which is a measure of the spread of a distribution is also very large with a large standard

error. This suggests that there could be large variability in the data. Moreover, the estimate for the variance of the frailty distribution (the heterogeneity parameter, θ) is 15.0981 in the best fitting model (calendar time frailty model with Weibull baseline hazard), and generally also very large in the other models. This implies that there is high between cluster (patient) variability and high correlation between events within same patient with a Kendall's tau estimated as $\hat{\theta}/(\hat{\theta}+2)=15.0981/(15.0981+2) = 0.8830$ with a standard error estimated as $2*S.E(\hat{\theta})/(2+\hat{\theta})^2 = (2*8.1367)/(2+15.0981)^2 = 0.0557$.

A further exploratory investigation was done by fitting the covariates one in turn in the model and it was observed that the estimated value of the scale parameter (λ) drastically reduced in all scenarios with a corresponding reduction in its standard error (results not shown). This could suggest the high estimate for λ observed in Tables 3.2.1 and 3.2.2 with all main effects is contributed by the fact that there were many parameters to estimate with a small sample size (events) compared to the reduced parameters scenario (single covariate models) where parameter estimation was more precise. Secondly, to check the influence of the scale parameter estimation on the model fit, the parameter was fixed at variable values and the change in the log likelihood observed (results not shown) and indeed model fit improved with an increase in the specified value of λ . This suggested an influence of the λ parameter estimation on model fit.

6. Discussion

The main objective of this study was to study the relationship between baseline patient characteristics and the risk of rehospitalisation in CHF patients. Being a clustered survival data setting with some time ordinality in repeated events, the popular one parameter gamma shared frailty model with calendar and gap time representations was used. In both cases, the Weibull and the exponential distributions were assumed for the baseline hazard functions. Of the 80 patients, 16 had rehospitalisations by the end of the study, with one patient having three rehospitalisations, and two patients having two rehospitalisations. The results showed that the frailty model with the Weibull baseline hazard provided a better fit to the data as compared to the frailty model with an exponential baseline hazard. For the time representations, the frailty model with the calendar time representation provided better fit as compared to the frailty model with the gap time representation, assuming Weibull baseline

hazards for both. The patient's heart rhythm status was found to be associated with the risk of rehospitalisation in CHF patients, with abnormal heart rhythm patients at a higher risk of being rehospitalised compared to if he/she had a normal heart rhythm. These results are in line with the findings of Korte *et al.* (2000) and Linssen *et al.* (2011) who also found abnormal heart rhythm patients to be at higher risk of being rehospitalised. Due to the sample size limitations in this study, it is important to note that these results may not rule out the possible prognostic ability of other covariates. Such covariates may include the N-terminal pro-brain natriuretic peptide, left ventricular ejection fraction, etc., that have been noted as well-known prognostic factors in chronic heart failure (Yukiko *et al.*, 2007). Based on the suggestion of Peduzzi *et al.* (1995), at least 10 events were required for each covariate to be investigated. Therefore, having just 20 events and even then, some clustered, estimating as much as six covariates poses a possibility of false negative/positive results (Peduzzi *et al.*, 1995; Bradburn *et al.*, 2003).

7. Conclusion

From our results, the Weibull baseline hazard with the calendar time representation provided better fit as compared to the exponential baseline hazard and the gap time representations respectively. Furthermore, given the data at hand, the patient's heart rhythm status was prognostic for time to rehospitalisation in telemonitored CHF patients with abnormal heart rhythm patients at higher risk of rehospitalisation. However, the reliability of these results may be questionable due to sample size limitations.

8. Limitations and Recommendations.

The main limitation to this study was that of small sample size (few patients that experienced the events). With this, one cannot completely rule out possible bias on parameter estimation and the fact that some potential prognostic factors could have been missed. Therefore, it is strongly recommended that this investigation be done in a larger study to be able to answer the research question.

9. References

1. Abrahantes, J. C., Legrand C., Burzykowski, T., Janssen, P., Ducrocq, V., Duchateau, L. (2007). Comparison of different estimation procedures for proportional hazards model with random effects. *Computational Statistics & Data Analysis*, **51**, 3913–3930.
2. American Heart Associations: *Classes of HeartFailure*, http://www.heart.org/HEARTORG/Conditions/HeartFailure/AboutHeartFailure/Classes-of-Heartfailure_UCM_306328_Article.jsp# (2011, Accessed 30th December 2013).
3. Anderson, G. F., Steinberg, E. P. (1985). Predicting hospital readmissions in the Medicare population. *Inquiry*, **22**, 251-258.
4. Bettencourt, P., Azevedo, A., Pimenta, J., Friões, F., Ferreira, S., Ferreira, A. (2004). N-Terminal–Pro-Brain Natriuretic Peptide Predicts Outcome After Hospital Discharge in Heart Failure Patients. *Circulation*, **110**, 2168-2174
5. Bradburn, M. J., Clark, T. G., Love, S. B., and Altman, D. G. (2003). Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. *British Journal of Cancer*, **89**, 605 – 611.
6. Cheng, V., Kazanagra, R., Garcia, A., Lenert, L., Krishnaswamy, P., Gardetto, N., Clopton, P., Maise, I A., (2001). A rapid bedside test for B-type peptide predicts treatment outcomes in patients admitted for decompensated heart failure: a pilot study. *Journal of the American college of Cardiology*, **37**, 386-391.
7. Clayton, D. G. (1978). A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incident. *Biometrika*, **65**, 141-151.
8. Collett, D. (1994). *Strategy for model selection. Modelling survival data in medical Research*. London: Chapman and Hall.
9. Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society*, **34**, 187–220.
10. Dendale, P., De Keulenaer, G., Troisfontaines, P., Weytjens, C., Mullens, W., Elegeert, I., Ector, B., Houbrechts, M., Willekens, K. and Hansen, D. (2011). Effect of a telemonitoring-

facilitated collaboration between general practitioner and heart failure clinic on mortality and rehospitalization rates in severe heart failure: the TEMA-HF 1 (TElemonitoring in the MAnagement of Heart Failure) study. *European Journal of Heart Failure*, **14**, 333-340.

11. Desai, M. M., Stauffer B.D., Feringa, H. H. H., Schreiner, G. C. (2009). Statistical Models and Patient Predictors of Readmission for Acute Myocardial Infarction *Circulation: Cardiovascular Quality and Outcomes*, **2**, 500-507.

12. Dokainish, H., Zoghbi, W. A., Lakkis, N. M., Ambriz, E., Patel, R., Quinones, M. A., Nagueh, S. F. (2005). Incremental predictive power of B-type natriuretic peptide and tissue Doppler echocardiography in the prognosis of patients with congestive heart failure. *Journal of the American college of Cardiology*. **45**, 1223-1226.

13. Douglas, L. M, Michael, R. B. (2005). Mechanisms and Models in Heart Failure: The Biomechanical Model and Beyond. *Circulation*, **111**, 2837-2849.

14. Duchateau, L. and Janssen, P. *The Frailty Model*. New York: Springer, 2008.

15. Duchateau, L., Janssen, P., Kezic, I. and Fortpied, C. (2003). Evolution of recurrent asthma event rate over time in frailty models. *Applied Statistics*, **52**, 355–363.

16. Formiga, F., Chivite, D., Manito, N., Osma, V., Miravet, S., Pujol, R. (2004). One-year follow-up of heart failure patients after their first admission, *QJM: An International Journal of Medicine* ,**97**, 81-86.

17. Frees, E. W., Carriere, J., Valdez, E. (1996). Annuity valuation with dependent mortality. *Journal of Risk Insurance*, **63**, 229-261.

18. Gackowski, A., Isnard, R., Golmard, J. L., Pousset, F., Carayon, A., Montalescot, G., Hulot, J. S., Thomas, D., Piwowarska, W., Komajda, M. (2004). Comparison of echocardiography and plasma B-type natriuretic peptide for monitoring the response to treatment in acute heart failure. *European Heart Journal*, **25**, 1788-1796.

19. Gheorghiade, M., Vaduganathan, M., Fonarow, G., Bonow, R. (2013). Rehospitalization for heart failure: problems and perspectives. *Journal of the American college of Cardiology*, **61**, 391-403.

20. Glidden, D. V., Vittinghoff, E. (2004). Modelling clustered survival data from multicentre clinical trials. *Statistics In Medicine*, **23**, 369–388.

21. Goonewardena, S. N., Gemignani, A., Ronan, A., Vasaiwala, S., Blair, J., Brennan, J. M., Shah, D. P., Spencer, K. T. (2008). Comparison of hand-carried ultrasound assessment of the inferior vena cava and N-terminal pro-brain natriuretic peptide for predicting readmission after hospitalization for acute decompensated heart failure. *Journal of the American college of Cardiology, Cardiovascular Imaging*, **1**, 595-601.
22. Hammill, B. G., Curtis, L. H., Fonarow, G. C., Heidenreich, P. A., Yancy, C. W., Peterson, E. D., Hernandez, A. F. (2011). Incremental Value of Clinical Data Beyond Claims Data in Predicting 30-Day Outcomes After Heart Failure Hospitalization. *Circulation, Cardiovascular Quality and Outcome*, **4**, 60-67.
23. Hanagal, D. D, Dabad, A. D. A. (2013.) Comparative Study of Shared Frailty Models for Kidney Infection Data with Generalized Exponential Baseline Distribution, *Journal of Data Science* ,**11**, 109-142.
24. Kiernan, K., Tao, J., Gibbs, P. (2012). Tips and Strategies for Mixed Modeling with SAS/STAT® Procedures URL <http://support.sas.com/resources/papers/proceedings12/332-2012.pdf>
25. Klein, J. P. and Moeschberger, M. L. *Survival analysis techniques for censored and truncated data*. New York: Springer, 2003.
26. Korte ,T., Jung, W., Ostermann, G., Wolpert, C., Spehl, S., Esmailzadeh, B. and Lüderitz, B. (2000). Hospital readmission after transvenous cardioverter/defibrillator implantation. A single centre study. *European Heart Journal*, **21**, **1186**-1191.
27. Linssen, G. C. M., Rienstra, M., Jaarsma, T., Voors, A. A., Gelder, I. C., Hillege, H. L., Van Veldhuisen, D. J. (2011). Clinical and Prognostic Effects of Atrial Fibrillation in Heart Failure Patients With Reduced and Preserved Left Ventricular Ejection Fraction. *European Journal of Heart Failure*, **13**, 1111-1120.
28. Logeart, D., Thabut, G., Jourdain, P., Chavelas, C., Beyne, P., Beauvais, F., Bouvier, E., Solal, A. C. (2004). Predischarge B-type natriuretic peptide assay for identifying patients at high risk of re-admission after decompensated heart failure. *Journal of the American college of Cardiology*, **43**, 635-641.
29. López, J. (2011). The heart failure epidemic. *Medicographia*, **33**, 363-369.

30. Murakami, M., Niwano, S., Koitabashi, T., Inomata, T., Satoh, A., Kishihara, J., Ishikawa S., Aoyama, Y., Niwano, H., Izumi, T. (2012). Evaluation of the impact of atrial fibrillation on rehospitalisation events in heart failure patients in recent years. *Journal of Cardiology*, **60**,36-41.
31. McGilchrist, C. A. and Aisbett, C.W. (1991). Regression with frailty in survival analysis. *Biometrics*, **47**, 461–466.
32. Njagi, E. N., Rizopoulos, D., Molenberghs, G., Dendale, P., and Willekens, K. (2013a). A joint survival-longitudinal modelling approach for the dynamic prediction of rehospitalization in telemonitored chronic heart failure patients. *Statistical Modelling*, **13**, 179-198.
33. Njagi, E.N., Molenberghs, G., Rizopoulos, D., Verbeke, G., Kenward, M.G., Dendale, P., and Willekens, K. (2013b). A flexible joint-modeling framework for longitudinal and time-to-event data with overdispersion. *Statistical Methods in Medical Research*. Published online before print July 18, 2013, doi: 10.1177/0962280213495994.
34. Nishio, Y., Sato, Y., Taniguchi, R., Shizuta, S., Doi, T., Morimoto, T., Kimura, T., Kita, T. (2007). Cardiac Troponin T vs Other Biochemical Markers in Patients With Congestive Heart Failure. *Circulation*, **71**, 631 –635.
35. Packer, M. (1992). The neurohormonal hypothesis: a theory to explain the mechanism of disease progression in heart failure. *Journal of the American college of Cardiology*, **20**, 248-254.
36. Peduzzi, P., Concato, J., Feinstein, A. R., Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression analysis II: accuracy and precision of regression estimates. *Journal of Clinical Epidemiology*, **48**, 1503–1510.
37. Sahu, S. K., Dey, D. K. (2000). A comparison of frailty and other models for bivariate survival data. *Lifetime Data Analysis*, **6**, 207-228.
38. Smith, D. M., Giobbie-Hurder, A., Weinberger, M., Oddone, E. Z., Henderson, W. G., Asch, D. A., Ashton, C. M., Feussner, J. R., Ginier, P., Huey, J. M., Hynes, D. M., Loo , L., Mengel, C. E. (2000). Predicting non-elective hospital readmissions: A multi-site study. *Journal of Clinical Epidemiology*, **53**, 1113–1118.

39. Thomas, J. W. (1996). Does risk-adjusted readmission rate provide valid information on hospital quality?. *Inquiry*, **33**, 258-270.
40. Vaupel, J. W., Manton, K. G., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16**, 439-454.
41. Yamokoski, L.M., Hasselblad, V., Moser, D.K., Binanay, C., Conway, G.A., Glotzer, J.M., Hartman, K.A., Stevenson, L.W., Leier, C.V. (2007) Prediction of rehospitalisation and death in severe heart failure by physicians and nurses of the ESCAPE trial. *Journal of Cardiac Failure*, **13**, 8-13.

10. Appendix

10.1. Appendix A

Derivation of Marginal density Functions

The assumed distribution for random frailty is the one parameter

$$\text{gamma given by: } f_{\varphi_i}(\varphi) = \frac{\varphi^{1/\theta-1} \exp(-\varphi/\theta)}{\theta^{1/\theta} \Gamma(1/\theta)},$$

with Γ the gamma function,.We note that $E(\varphi) = 1$ and $Var(\varphi) = \theta$

Frailty Model with the weibull distributed baseline hazard($h_o(t) = \lambda \rho t^{\rho-1}$)

Gap time representation

The conditional Hazard function for the j^{th} event of subject i , given the frailty term is given by :

$$h_{ij}(t/\varphi) = h_o((t - y_{ij1})) \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}),$$

where $h_o(t)$ is the baseline hazard , φ_i is called the frailty for the i^{th} subject, $\boldsymbol{\beta}$ is the fixed effect vector of dimension p , \mathbf{x}_{ij} is the vector of covariates.

Let $Z = \text{Riskend}(y_{ij2}) - \text{Riskstart}(y_{ij1})$ (the time interval from the risk of an event to when the event occur).

Therefore the Cumulative hazard function for at risk period j of subject i is given by:

$$H_{ij}(t) = \int_0^{\text{riskend}-\text{riskstart}} h_{ij}(t) dt = \int_0^Z h_{ij}(t) dt .$$

The Conditional survivor density for subject i , given frailty term:

$$f(t/\varphi) = h_{ij}(t/\varphi_i) s(t) \quad \text{Where } s(t) \text{ is the survival function}$$

$$= [\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_0^z h_{ij}(t) dt\right); \text{ where } \delta_{ij} \text{ is the censoring indicator}$$

$$\begin{aligned}
&= [\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_0^z \lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \\
&= [\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \rho \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) \int_0^Z t^{\rho-1} dt) \\
&= [\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho)
\end{aligned}$$

It follows therefore that, the marginal density for the i^{th} subject is given by:

$$\begin{aligned}
f(t) &= \int_0^\infty [\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \\
&= \frac{[\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^\infty \varphi_i^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho) \varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta) d\varphi_i \\
&= \frac{[\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^\infty \varphi_i^{(1/\theta+\delta_{ij})-1} \exp(-\varphi \lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho + 1/\theta) d\varphi_i
\end{aligned}$$

Which reduces to :

$$\begin{aligned}
&\frac{[\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(1/(\lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho + 1/\theta))^{-(1/\theta+\delta_{ij})}} \\
&= \frac{[\lambda \rho (t - y_{ij1})^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z^\rho + 1/\theta)^{(1/\theta+\delta_{ij})}}
\end{aligned}$$

Where Γ is the gamma function.

Calendar time representation

For the calendar time representation, we are interested in exact times when the risk begins and when the event occurs.

So we let $y_{ij1} = Riskstart$ and $y_{ij2} = riskend$

Hence the conditional density becomes

$$\begin{aligned}
f(t/\varphi) &= [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}} \exp\left(-\int_{y_{ij1}}^{y_{ij2}} h_{ij}(t) dt\right) \\
&= [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}} \exp\left(-\int_{y_{ij1}}^{y_{ij2}} \lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \\
&= [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}} \exp\left(-\lambda \rho \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) \int_{y_{ij1}}^{y_{ij2}} t^{\rho-1} dt\right) \\
&= [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}} \exp\left(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) (y_{ij2}^\rho - y_{ij1}^\rho)\right)
\end{aligned}$$

And the Marginal density for the i^{th} subject is therefore given by

$$\begin{aligned}
f(t) &= \int_0^\infty [\lambda \rho t^{\rho-1} \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}} \exp\left(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) (y_{ij2}^\rho - y_{ij1}^\rho)\right) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \\
&= \frac{[\lambda \rho t^{\rho-1} \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^\infty \varphi_i^{s_{ij}} \exp\left(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) (y_{ij2}^\rho - y_{ij1}^\rho)\right) \varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta) d\varphi_i \\
&= \frac{[\lambda \rho t^{\rho-1} \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^\infty \varphi_i^{(1/\theta + s_{ij})-1} \exp\left(-\varphi \lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) (y_{ij2}^\rho - y_{ij1}^\rho) + 1/\theta\right) d\varphi_i
\end{aligned}$$

Which reduces to :

$$\frac{[\lambda \rho t^{\rho-1} \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{s_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + s_{ij})}{(\lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) (y_{ij2}^\rho - y_{ij1}^\rho) + 1/\theta)^{(1/\theta + s_{ij})}}$$

where Γ is the gamma function.

Exponential distribution with baseline hazard : $h_o(t) = \lambda$

Gap time representation

Again we are not interested in the exact times when the risk for an even starts and ends, but just the time interval when the risk for an event starts to when the event occurs.

The conditional hazard function, given the frailty is given by

$$\begin{aligned}
h_{ij}(t/\varphi) &= h_o(t)\varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) \\
&= \lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})
\end{aligned}$$

The conditional density function for subject i , given the frailty, is therefore:

$$f(t/\varphi) = [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_0^z h_{ij}(t) dt\right)$$

where δ_{ij} is the censoring indicator and Z is the

time interval from when the risk for an event starts to when it ends.

$$\begin{aligned}
&= [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_0^z \lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \\
&= [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z)
\end{aligned}$$

It again follows that the marginal density for the i^{th} subject is given by:

$$\begin{aligned}
f(t) &= \int_0^\infty [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \\
&= \frac{[\lambda \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^\infty \varphi_i^{\delta_{ij}} \exp(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z) \varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta) d\varphi_i \\
&= \frac{[\lambda \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^\infty \varphi_i^{(1/\theta+\delta_{ij})-1} \exp(-\varphi \lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) Z + 1/\theta) d\varphi_i
\end{aligned}$$

Which again reduces to:

$$= \frac{[\lambda \exp(\boldsymbol{\beta} \mathbf{x})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda_i \exp(\boldsymbol{\beta} \mathbf{x}) Z + 1/\theta)^{(1/\theta+\delta_{ij})}}$$

Where Γ is the gamma function.

Calendar time representation

Under The calendar time representation, the derivation flow in same way as in gap time representation to conditional density function which becomes :

$$\begin{aligned}
f(t/\varphi) &= [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_{t_1}^{t_2} h_{ij}(t) dt\right) \\
&= [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\int_{t_1}^{t_2} \lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) dt\right) \\
&= [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta}) \int_{t_1}^{t_2} dt\right) \\
&= [\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}} \exp\left(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})(y_{ij2} - y_{ij1})\right)
\end{aligned}$$

And the Marginal density for the i^{th} subject is given by:

$$\begin{aligned}
f(t) &= \int_0^{\infty} [\lambda \varphi_i \exp(\boldsymbol{\beta} \mathbf{x})]^{\delta_{ij}} \exp\left(-\lambda \varphi_i \exp(\boldsymbol{\beta} \mathbf{x})(y_{ij2} - y_{ij1})\right) \frac{\varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta)}{\theta^{1/\theta} \Gamma(1/\theta)} d\varphi_i \\
&= \frac{[\lambda \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^{\infty} \varphi_i^{\delta_{ij}} \exp\left(-\lambda \varphi_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})(y_{ij2} - y_{ij1})\right) \varphi_i^{1/\theta-1} \exp(-\varphi_i / \theta) d\varphi_i \\
&= \frac{[\lambda \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \int_0^{\infty} \varphi_i^{(1/\theta+\delta_{ij})-1} \exp\left(-\varphi_i \lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})(y_{ij2} - y_{ij1}) + 1/\theta\right) d\varphi_i,
\end{aligned}$$

which reduces to $\frac{[\lambda \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})]^{\delta_{ij}}}{\theta^{1/\theta} \Gamma(1/\theta)} \times \frac{\Gamma(1/\theta + \delta_{ij})}{(\lambda_i \exp(\mathbf{x}_{ij}^t \boldsymbol{\beta})(y_{ij2} - y_{ij1}) + 1/\theta)^{(1/\theta+\delta_{ij})}}$;

where Γ is again the gamma function.

10.2 Appendix B.

Selected Codes

R codes

```
/* Kaplan-Meier Plot*/  
  
/*Dataset name=data3*/  
library(survival)  
msurv <- with(data3, Surv(day, status==1))  
mfit <- survfit(Surv(day, status == 1)~ 1, data = data3)  
plot(mfit, ylab="Survival Probability",xlab="Time to Rehospitalisation(Days)")
```

SAS Codes

```
/*Dataset name= thesis.*/  
/*Delete observations with missing NTproBNP measurements*/  
  
data reduced;  
  set thesis;  
  if NTproBNP= . then delete; /  
run;  
  
/*rescaling variables*/  
data sreduced;  
set reduced;  
sage= age/(95);  
sNYHA= nyha/4;  
sLVEF= LVEF/80;  
sNTproBNP= NTproBNP/37690;  
  
cresp=day/185;  
sday=cresp;/*cresp=scaled Response to use in calender time representation */  
sriskstart=riskstart/185;  
sriskend=riskend/185;  
sresp=sriskend-sriskstart;/*scaled Response to use in Gap time representation*/  
run;quit;  
  
/*Programming categorical variables to use in prog lifereg, to get starting values*/  
  
/*Lifereg models "0s" in categorical variables while Nlmixed models "1s".  
So we now reverse the coding in sex and heartrym to use in lifereg for comparability with  
NLMIXED */  
data treduced;  
set sreduced;  
if sex= 1 then tsex=0;  
else tsex=1;  
if heartrym= 1 then theartrym= 0;  
else theartrym=1;run;quit;
```

```

/*WEIBULL BASELINE */
/*Get Starting Values*/

proc lifereg data= treduced ;
class tsex theartrym;
model sday*status(0)= sage tsex sNYHA sLVEF sNTproBNP theartrym ;
run;quit;

/*Starting Values Gotten*/
/*
parms beta1=-0.737325248 beta2=-0.365167593 beta3=-0.404244568 beta4=-0.264780192
beta5=1.473892538 beta6=0.793245747
lamda=0.43397644
rho=0.842176183;*/

/*Gap time Frailty Model*/

proc nlmixed data= sreduced tech=newrap;
bounds lamda > 0, rho > 0,theta > 0;
parms beta1=-0.737325248 beta2=-0.365167593 beta3=-0.404244568 beta4=-0.264780192
beta5=1.473892538 beta6=0.793245747
lamda=0.43397644
rho=0.842176183;
mu=exp(beta1*sage + beta2*sex + beta3*sNYHA + beta4*sLVEF + beta5*sNTproBNP +
beta6*heartrym );

ll= log((lamda*rho*((sriskend-sriskstart)**(rho-1))*mu)**status)
+ lgamma((1/theta) +status)
-log(theta**(1/theta))
-lgamma(1/theta)
-log((lamda*mu*((sriskend-sriskstart)**rho) +(1/theta))**((1/theta) + status));

model sresp ~ general(ll);
run; quit;

/*Calender time Frailty Model*/
proc nlmixed data= sreduced tech=newrap ;
bounds lamda > 0, rho > 0,theta > 0;
parms beta1=-0.737325248 beta2=-0.365167593 beta3=-0.404244568 beta4=-0.264780192
beta5=1.473892538 beta6=0.793245747
lamda=0.43397644
rho=0.842176183;

mu=exp(beta1*sage + beta2*sex + beta3*sNYHA + beta4*sLVEF + beta5*sNTproBNP +
beta6*heartrym );

ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
-log(theta**(1/theta))

```

```

        -lgamma(1/theta)
        -log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta))**((1/theta)+status));
model cresp ~ general(ll);
run; quit;

/*EXPONENTIAL BASELINE*/

/*Get Starting Values*/

proc lifereg data= treduced ;
class tsex theartrym;
model sday*status(0)= sage tsex sNYHA sLVEF sNTproBNP theartrym/Distribution=
Exponential;
run;quit;

/*Starting Values Gotten*/
/*
parms beta1=-0.7563 beta2=-0.3678 beta3=-0.3697 beta4=-0.2562 beta5=1.475
beta6=0.7901
lamda=0.436310994 */

/*Gap time Frailty Model*/

proc nlmixed data= sreduced tech=newrap;
bounds lamda > 0,theta > 0;
parms beta1=-0.7563 beta2=-0.3678 beta3=-0.3697 beta4=-0.2562 beta5=1.475
beta6=0.7901
lamda=0.436310994;

mu=exp( beta1*sage + beta2*sex + beta3*sNYHA + beta4*sLVEF + beta5*sNTproBNP +
beta6*heartrym);

LL = status*log(lamda*mu)+lgamma((1/theta) +status) - log(theta**(1/theta)) -
lgamma(1/theta)
-((1/theta) + status)*log(lamda*mu*(sriskend-sriskstart) + (1/theta));

model sresp ~ general(ll);
run; quit;

/*Calender time Frailty Model*/
proc nlmixed data= sreduced tech=newrap;
bounds lamda > 0,theta > 0;
parms beta1=-0.7563 beta2=-0.3678 beta3=-0.3697 beta4=-0.2562 beta5=1.475
beta6=0.7901
lamda=0.436310994 ;

mu=exp( beta1*sage + beta2*sex + beta3*sNYHA + beta4*sLVEF + beta5*sNTproBNP +
beta6*heartrym);

```

```
ll= status*log(lamda*mu) +lgamma((1/theta) + status) -log(theta**(1/theta))-lgamma(1/theta)
  -((1/theta) +status)*log(lamda*mu*(sriskend-sriskstart) + (1/theta));
```

```
model cresp ~ general(ll);
```

```
run; quit;
```

```
/*Univariate Analysis Using Calender time Frailty Model with Weibull baseline(Best Fitting Model)*/
```

```
/*AGE*/
```

```
/*Get starting Values*/
```

```
proc lifereg data= tduced ;
```

```
model sday*status(0)= sage ;
```

```
run;quit;
```

```
/*Starting Values Gotten*/
```

```
/*
```

```
parms beta1=0.53023722
```

```
lamda=0.160906874
```

```
rho=0.835282325
```

```
*/
```

```
/*Frailty Model*/
```

```
proc nlmixed data= sreduced tech=newrap ;
```

```
bounds lamda > 0, rho > 0,theta > 0;
```

```
parms beta1=0.53023722
```

```
lamda=0.160906874
```

```
rho=0.835282325;
```

```
mu=exp(beta1*sage );
```

```
ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
```

```
-log(theta**(1/theta))
```

```
-lgamma(1/theta)
```

```
-log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta))**((1/theta)+status));
```

```
model cresp ~ general(ll);
```

```
run; quit;
```

```
/*SEX*/
```

```
/*Get starting Values*/
```

```
proc lifereg data= tduced ;
```

```
model sday*status(0)= tsex ;
```

```
run;quit;
```

```
/*Starting Values Gotten*/
```

```
/*
```

```
parms beta2=0.197523634
```

```
lamda=0.227246472
```

```
rho=0.836610056 */
```

```
/*Frailty Model*/
```

```
proc nlmixed data= sreduced tech=newrap ;
```

```
bounds lamda > 0, rho > 0,theta > 0;
```

```

parms beta2=0.197523634
      lamda=0.227246472
      rho=0.836610056;

mu=exp(beta2*sex);

ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
      -log(theta**(1/theta))
      -lgamma(1/theta)
      -log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta)**((1/theta)+status)));
model cresp ~ general(ll);
run; quit;

/*NYHA*/
/*Get starting Values*/
proc lifereg data= treduced ;
model sday*status(0)= sNYHA ;
run;quit;
/*Starting Values Gotten*/
/*
parms beta3=1.338390612
      lamda=0.087396423
      rho=0.838222967
      */
/*Frailty Model*/
proc nlmixed data= sreduced tech=newwrap ;
bounds lamda > 0, rho > 0,theta > 0;
parms beta3=1.338390612
      lamda=0.087396423
      rho=0.838222967;

mu=exp(beta3*sNYHA);

ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
      -log(theta**(1/theta))
      -lgamma(1/theta)
      -log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta)**((1/theta)+status)));
model cresp ~ general(ll);
run; quit;

/*LVEF*/

/*Get starting Values*/
proc lifereg data= treduced ;
model sday*status(0)= sLVEF ;
run;quit;
/*Starting Values Gotten*/
/*
parms beta4=-0.31937829

```

```

lamda=0.282977403
rho=0.835631319
*/
/*Frailty Model*/
proc nlmixed data= sreduced tech=newrap ;
bounds lamda > 0, rho > 0,theta > 0;
parms beta4=-0.31937829
      lamda=0.282977403
      rho=0.835631319;

mu=exp(beta4*sLVEF);

ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
     -log(theta**(1/theta))
     -lgamma(1/theta)
     -log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta))**((1/theta)+status));
model cresp ~ general(ll);
run; quit;

/*NTproBNP*/

/*Get starting Values*/
proc lifereg data= treduced ;
model sday*status(0)= sNTproBNP ;
run;quit;
/*Starting Values Gotten*/
/*
parms beta5=1.079868985
      lamda=0.20745179
      rho=0.83984211
      */

/*Frailty Model*/
proc nlmixed data= sreduced tech=newrap ;
bounds lamda > 0, rho > 0,theta > 0;
parms beta5=1.079868985
      lamda=0.20745179
      rho=0.83984211;

mu=exp(beta5*sNTproBNP);

ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
     -log(theta**(1/theta))
     -lgamma(1/theta)
     -log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta))**((1/theta)+status));
model cresp ~ general(ll);
run; quit;

```

```

/*Heartrym*/

/*Get starting Values*/
proc lifereg data= tReduced ;
model sday*status(0)= tHeartrym ;
run;quit;
/*Starting Values Gotten*/
/*
parms beta6=-0.665128764
      lamda=0.331425026
      rho=0.838855801
      */

/*Frailty Model*/
proc nlmixed data= sReduced tech=newwrap ;
bounds lamda > 0, rho > 0, theta > 0;
parms beta6=-0.665128764
      lamda=0.331425026
      rho=0.838855801;

mu=exp(beta6*Heartrym);

ll= log((lamda*rho*(sday**(rho-1))*mu)**status) + lgamma((1/theta)+status)
    -log(theta**(1/theta))
    -lgamma(1/theta)
    -log((lamda*mu*(sriskend**rho - sriskstart**rho) + (1/theta))**((1/theta)+status));
model cresp ~ general(ll);
run; quit;

```

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:

Survival modelling of rehospitalization in telemonitored chronic heart failure patients

Richting: **Master of Statistics-Biostatistics**

Jaar: **2014**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

Njilefac, Richard

Datum: **6/02/2014**