Ilorin Journal of Science
Volume 1, Number 1, 2014, pp. 61 – 73 (Printed in Nigeria)
ISSN: 2408 – 4840 © 2014 Faculty of Physical Sciences, University of Ilorin
https://doi.org/10.54908/iljs.2014.01.01.005

ILORIN
JOURNAL OF SCEINCE

ILJS-14-001

An Application of Semiparametric Structured Additive Model to Cancer Data

Abiodun, A. A.

Department of Statistics, University of Ilorin, P.M.B. 1515 Ilorin, Ilorin, Nigeria

Abstract

In many epidemiological studies where times to event data are clustered, introducing frailties in the Cox model can account for heterogeneity induced by such clustering. Analyses were carried out using data collected on a sample of cancer patients from University of Ilorin Teaching Hospital, using Full Bayesian inference based on Markov Chain Monte Carlo (MCMC) simulation technique. The approach allows the estimation of very complex and realistic models. The results showed that sex and age were significant risk factors associated with death from some selected cancers. The risk of dying from the selected cancers was observed to progressively increase as age of patient's increases. Using Deviance Information Criterion (DIC) for model comparison, it was observed that the flexible semi parametric additive P-spines model, which allows for nonlinearity due to metrical covariate age, was better than the model that introduced metrical age linearly as fixed effect. It was also found that models that accounted for heterogeneity induced by clustering observations were more adequate than those that ignored it. On effect of interaction between sex and age on the death due to cancer, the model that contained interaction between sex and age when metrical age was modeled nonlinearly was observed to be better than those that modeled metrical or categorized age as linear effect.

Keywords: Survival time, Censoring, Independence, Frailty, Markov Chain Monte Carlo

1. Introduction

Survival time data are usually non-normal due to a number of censored observations thereby making analysis of such skewed data different from standard analyses. The most common censoring in survival time data is right censoring, which occurs when the actual time a subject experiences the event of interest is not known. In this type of censoring, it is assumed for some individuals in the study that there is a time to event T^*, and the right censoring time C, where the T^*'s are assumed to be independently and identically distributed with density function $f(t)$ and survival function $S(t)$. The exact survival time T of any individual will be known if and only if $T^* \leq C$. If however, $T^* > C$, then the individual is a survivor and the exact survival time is censored at C. Thus the observed time is $T = \min(T^*, C)$ and the data for such

Corresponding Author: Abiodun, A. A.
Email: alfredabiodun1@gmail.com

a design can be represented by a pair of random variables (T,δ), where δ indicates whether the survival time T corresponds to an event (δ =1) or is right censored (δ =0).

Often times, survival data contains tied observations, and these need be taken care of during analysis. The ideal method of handling ties is the Exact method of partial likelihood under Cox proportional hazard model formulation. This is however computer intensive (Huang and Liu, 2007). The methods by Breslow (1974) and Efron (1977) are much simpler. Adeleke *et al*. (2013) applied these estimation methods to breast cancer data under non-proportional parametric and semiparametric survival models.

Often times, survival data contains tied observations and the estimation methods for handling such data include Exact method of partial likelihood by Breslow (1974) and Efron (1977) methods under Cox proportional hazard model formulation,are much simpler. Adeleke *et al.* (2013) analyzed data on cancer patients using these methods.

In several survival studies, failure times are aggregated in clusters and in such situation, subjects belonging to the same cluster are similar with respect to certain characteristics so that their survival times are correlated whereas the survival times of subjects belonging to different clusters are independent.

A standard statistical approach to model clustered failure time data is the frailty model, which is a random effects model for survival data, where frailties are usually introduced as multiplicative random terms in the proportional hazards model (Cox, 1972). Observations belonging to the same cluster share the same characteristics than those belonging to different clusters. Frailty model was proposed by Clayton (1978), where the dependence between subjects was modeled by a common random effect (shared frailty model). This was extended by the correlated frailty model (Yashin *et al*., 1995; Xue and Brookmeyer, 1996; Yau and McGilchrist, 1997; Ripatti and Palmgren, 2000). Commonly used frailty distributions are the gamma, normal and lognormal distributions. Shared positive stable frailty model has also attracted attentions recently (Liu, et al, 2011). This is documented in Fine,Glidden and Lee (2003) and also in Martinussen and Pipper (2005). Liu et al (2011) proposed a covariate-dependent positive stable shared frailty model under a unified framework, where the marginal regression parameters and the covariate effects on the frailty distribution can be consistently estimated. Frailty models with flexible distributions of the frailties have also been proposed in the Bayesian context (Kim and Dey, 2008; Kom´arek and Lesaffre, 2008, 2009; Callegaro and Iacobelli, 2012).

2. Materials and Methods

2.1 Cox Proportional Hazards Model with Shared Frailty

One popular regression model formulation that is often used in survival analysis is the Cox proportional hazards model (Cox, 1972). The model utilizes the hazard function λ(t), defined as the probability of experiencing event of failure in the small interval $(t, t+\Delta t)$, given that such an event has not been experienced prior to time *t*, and it is expressed as

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr\left(t \le T \le t + \Delta t \middle| T > t\right)}{\Delta t} \tag{1}$$

Suppose that we observe censored survival data from G clusters with n_g subjects in *g*th cluster ($g = 1, \ldots, G$ and $i = 1, \ldots, n_g$). Let $t_{ig} = \min(T_{ig}, C_{ig})$ be the observed time for *i*th subject in the *g*th cluster, where T_{ig} is the true survival time and C_{ig} is the censoring time, then the survival data can be given in the form $(t_{ig}, \delta_{ig}, z_{ig})$, where t_{ig}, is the time to failure of the *i*th subject in *g*th cluster, δ_{ig} is the censoring indicator such that for the δ_{ig} *i*th subject in cluster *g*, δ_{ig} =1 if event of failure occurs to the subject at time t_{ig},, and δ =0 if the time is right censored. If z_{ig} is the vector of covariates thought to be associated with t_{ig} and v_g is the cluster specific random effect (frailty) shared by all subjects in cluster *g*, then conditional on z_{ig} and v_g, the Cox proportional hazards model (Cox, 1972), extended to include the frailties $\exp(v_g)$, is given by

$$\lambda(t_{ig} | z_{ig}, v_g) = \lambda_0(t_{ig}) \exp(\gamma' z_{ig} + v_g), \tag{2}$$

where $\lambda_0(t_{ig})$ is an unspecified baseline hazard function, γ is a vector of the regression parameters and v_g , $g = 1, \ldots, G$ are independent and identically distributed frailties with density function $f(v)$.

The log-likelihood, conditional on the value of v_g can be expressed as

$$l_g(\gamma, \lambda_0, v_g) = \sum_{i}^{ng} [\lambda_{ig} \{\log \lambda_0(t_{ig}) + \gamma' z_{ig} + v_g\} - \Lambda_0(t_{ig}) \exp(\gamma' z_{ig}) + v_g], \tag{3}$$

where $\Lambda_0(t)$ denotes the cumulative baseline hazard function obtained by

$\Lambda_0(t) = \int_0^t \lambda_0(s) ds$. The augmented loglikelihood which can be obtained by adding log of the density of the frailties to (3) is given as

$$l_g^{(aug)}(\gamma, \lambda_0, v_g) = \sum_i^{n_g} [\lambda_{ig} \{\log \lambda_0(t_{ig}) + \gamma' z_{ig} + v_g\} - \Lambda_0(t_{ig}) \exp(\gamma' z_{ig}) + v_g] + \log f(v_g) \quad (4)$$

By integrating the likelihood function with respect to $f(v_g)$, we obtain marginal likelihood for *g*th cluster as

$$l_g^{(marg)}(\gamma, \lambda_0, \theta) = \int_{-\infty}^{\infty} \exp\left\{l_g^{aug}(\gamma, \lambda_0, v_g)\right\} dv_g \quad (5)$$

where θ denotes the parameter vector of the distribution $f(v)$.

The full loglikelihood is then given as

$$l(\gamma, \lambda_0, \theta) = \sum_{g=1}^{G} \log L_g^{(marg)}(\gamma, \lambda_0, \theta) \quad (6)$$

Model in (2) can be written as.

$$\lambda(t_{ig} | z_{ig}, v_g) = \exp\{\eta_i(t)\}, \quad (7)$$

with $\eta_i(t) = f_0(t) + \gamma' z_{ig} + b_g$,

where $f_0(t)=\log \lambda_0(t)$

An assumption often made in investigating the relationship between response variable and a set of covariates is that of linear effects of the metrical covariates on the response variable, such assumption on which the model expressed in (7) is based is too restrictive because in practical situations, effect of covariate such as age may be nonlinear and may not be adequately described by a linear relationship. Thus extending Hennerfeind *et al.* (2006), the predictor in (7) may be replaced with a more flexible semiparametric structured additive predictor that incorporates the complexity of nonlinearity in the same model framework. This is given by

$$\eta_i(t) = f_0(t) + f_j(x_{ij}) + \gamma' z_{ig} + b_g, \quad (8)$$

where

$f_0(t)=\log \lambda_0(t)$ is the log-baseline effect

f_j is the nonlinear effect of a covariate x_j

γ is the vector of the usual linear effects of categorical variables.

b_g is the cluster specific unstructured random effect (frailty) with $b_{ig}=b_g$ if *ith* subject is in cluster g, g =1,…,G. Clearly, b_g are usually assumed to be independent realizations from normal or log-gamma distribution with known mean and unknown variance.

2.2 Modeling Interaction Effect

Often, a model may have covariates z_1 and z_2 say, such that the effect of the two cannot be separated. In such a model, it will be inadequate to determine the incremental effects of both variables separately on the response variable. An interaction between two variables means that the effect of one variable on the outcome of interest is different depending on the level of the other variable (Wassertheil-Smoller, 2004). Interaction in Logistic and Cox regression models are inherently multiplicative, in which case the joint effect of the two variables is greater than the product of the individual effects of each variable. One way to describe such interaction effect is to add variable $z_3=z_1*z_2$ to the model (Harrell, 2001). In the field of biostatistics and epidemiology, some types of interactions that have consistently been found to be important in predicting outcomes include interactions between treatment and severity of diseases, between age and risk factors, between age and type of diseases.

2.3 Bayesian Inference

One commonly used inference method for frailty model is Full Bayesian analysis via Markov Chain Monte Carlo (MCMC) technique (Jones, 2004). In this method, each of the parameters in the model is iteratively re-sampled using its conditional densities given the current values of other parameters. For defining priors and developing posterior analysis, the predictor in (8) need be written in generic matrix notation. We thus express $f_0(t)$, f and b in matrix product of an appropriately defined designed matrix **Z** which leads to re-expressing (5) as

$$\boldsymbol{\eta = Z_0\beta_0(t)+Z_1\beta_1+\ldots.+Z_m\beta_m+V\gamma} \qquad (9)$$

Assignment of priors are as follows: For fixed effect parameter γ, diffused prior

$P(\gamma) \propto const\,\mathrm{P}(\gamma)$ is assumed

For non linear effects f and *b*, the general form of priors for β_j can be put in the form

$$p(\beta_j|\tau_j^2) \propto \exp\left(-\frac{1}{2\tau_j^2}\beta_j' K_j \beta_j\right),$$

where K_j is a precision or penalty matrix which shrinks parameters towards zero or penalizes too abrupt jumps between neighbouring parameters. It also depends on the prior assumptions about smoothness of f_j and the type of covariate. For example, for P-splines with first order random walk penalty, K_j is given by

$$K_j = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix},$$

and for an independent and identical random effect, the penalty matrix is the identity matrix, i.e. K_j =I. The variance parameter τ_j^2 controls the tradeoff between flexibility and smoothing and an inverse gamma prior (the conjugate prior) is assumed. i.e. $\tau_j^2 \sim IG(a,b)$.

For the baseline and non-linear effects ($g_0(t)$) and continuous covariate (f_j), Bayesian P-splines prior as in Lang and Brezger (2004) has been assigned and for the random effect (b_j), independently and identically distributed Gaussian Prior, $b_g \sim N(0, \tau_b^2)$ has been assigned.

Monte Carlo simulation methods are based on the principle of posterior distribution of sampling and subsequent use of these simulated samples for estimating the posterior distribution.

Suppose that $\beta = (\beta_0, ...\beta_m)'$ denote the vector of all regression coefficients in the generic notation for the functions, γ denotes the vector of linear effects and $\tau^2 = (\tau_o, \tau_m)$ which is the vector of all variance components, then full Bayesian inference is based on the posterior distribution $p(\beta, \gamma, \tau^2 | data) \propto L(\beta, \gamma, \tau^2) p(\beta, \gamma, \tau^2)$.

This is based on the assumptions that observations are independent conditional on covariates and entire set of parameters, and that prior distribution for fixed random and hyperparameters are mutually independent. The posterior distribution can therefore be written as

$$p(\beta, \gamma, \tau^2 | data) \propto L(\beta, \gamma, \tau^2) . \left\{ \prod_{j=0}^{m} p(\beta_j | p(\tau_j^2)) p(\tau_j^2) \right\} p(\gamma).$$

For updating the parameter vectors β_j which correspond to the functions $f_j(x_j)$, fixed linear effect γ and random effect b_g, Metropolis Hasting algorithm (Gamerman, 1997) and Brezger & Lang, 2006), based on iteratively weighted least squares (IWLS) proposals has been used.

3. Results and Discussion

Data were collected on a sample of 240 cancer patients who were admitted at the University of Ilorin Teaching Hospital (UILTH) from 2006 to 2012. The record of each patient contained information of variables length of stay in hospitals (in days), sex, age of patients and outcome which indicates whether the patient is dead or alive. Survival time is defined as length of admission before death occurs, while those who were still alive at the time of data collection were right-censored. Ten types of cancer were included in the study, excluding prostate and breast cancers because they are gender related and may possibly introduce gender bias into the analysis. Similar data have been analyzed in Abiodun (2009).

Observations on patients having similar cancer are expected to be more correlated than for patients with different types of cancer. This is because many different cancer/tumour types with distinct sites, present different clinical behaviours (Eschenbach & Collins, 2005). Therefore, cancer type has been included in the model as random effect (frailty)

Sex and metrical age were first fitted as linear effects. Thus, the fitted model, dropping subscript *i* is

$$\eta(t) = f_0(t) + sex\gamma_1 + age\gamma_2$$

To gain more insight into the analysis with respect to gender differentials, separate models were fitted for males and females. Also, since the assumption of linear effects of metrical covariates such as age on the predictor is too restrictive as discussed in section (2), two alternative approaches were considered (Abiodun, 2009). Firstly, age was grouped into three categories and modelled as linear effects with diffuse prior; secondly, it was incorporated additively in the predictor using smooth regression function and modelled as nonlinear effect using P-splines prior as in Lang and Brezger (2004). In this paper age was grouped into "less than 23 years" (reference group), "23-39 years", "40-55 years", and "greater than 55 years". The age grouping has demographic justification guided by number of observations in each group such that no group has too few observations that may render the results incredible. The research interest thus include: investigating the superiority of modelling metrical age, using P-splines over fitting it as linear effect, comparing models that included cancer types as random effects (frailty) with models that assumed independence (no frailty) for the survival times of the patients ignoring frailty and modelling interactions between sex and age for all patients. Model comparison was based on Deviance information criterion (DIC), introduced

by Spiegeihalter *et al* (2002), which is a Bayesian analogue of Akaike Information criterion (AIC). The following models were fitted for all patients and separately for males and females.

Models for all patients

Model 1: $\eta = f_0(t) + sex\gamma + age\gamma$ (fitting metrical age as linear effect)

Model 2: $\eta = f_0(t) + sex\gamma + age\gamma + b_g$ (Model 1 with frailty)

Model 3: $\eta = f_0(t) + sex\gamma + f_{age}$ (fitting metrical age with P-splines)

Model 4: $\eta = f_0(t) + sex\gamma + f_{age} + b_g$ (Model 3 with frailty)

Model 5: $\eta = f_0(t) + sex\gamma + age_1\gamma + age_2\gamma + age_3\gamma$ (fitting categorized age as linear effect)

Model 6: $\eta = f_0(t) + sex\gamma + age_1\gamma + age_2\gamma + age_3\gamma + b_g$ (Model 5 with frailty)

Model 7: $\eta = f_0(t) + sex\gamma + f_{age} + sex^* age\gamma + b_g$ (Model 2 with interaction effect of sex and metrical age)

Model 8: $\eta = f_0(t) + sex\gamma + f_{age} + sex^* f_{age} + b_g$ (Model 4 with interaction effect of sex and metrical age)

Model 9: $\eta = f_0(t) + sex\gamma + age_1\gamma + age_2\gamma + age_3\gamma + sex^* age_1\gamma + sex^* age_2\gamma + sex^* age_3\gamma + b_g$ (Model 6 with interaction effect of sex and categorical age)

Gender Models

Model 1: $\eta = f_0(t) + age\gamma$ (fitting metrical age with linear as linear effect)

Model 2: $\eta = f_0(t) + age\gamma + b_g$ (Model 1 with frailty)

Model 3: $\eta = f_0(t) + f_{age}$ (fitting continuous age with P-splines)

Model 4: $\eta = f_0(t) + f_{age} + b_g$ (Model 3 with frailty)

Model 5: $\eta = f_0(t) + age_1\gamma + age_2\gamma + age_3\gamma$ (fitting categorized age as linear effect)

Model 6: $\eta = f_0(t) + age_1\gamma + age_2\gamma + age_3\gamma + b_g$ (Model 5 with frailty)

Table 1 shows the hazard ratio, standard error and the 95% confidence intervals when sex and metrical age were fitted as linear effect (Model 1). As observed, effects of sex and age are seen significant (the confidence intervals do not include 1). Table 2 shows the results of

linear effects using categorized age for all, male and female patients. As observed in the table, age of the patients have significant influence on the death from cancer. For all patients, the hazard ratio for male patients is 0.465, meaning that male patients have lower risk of dying from cancer than their male counterparts. Age is also observed to have significant effect on mortality due to cancer. The hazard ratio of patients who are 23-39years old is 1.34, meaning that patients between age 23-39 years are 1.34 times more likely to die from cancer than those aged less than 23 years. The corresponding hazard ratio for male patients is 1.56, while it is 1.25 for female patients. Patients who are 40-55 years old have hazard ratios of 1.52, 1.705 and 1.63 for all patients, male and female patients respectively, while patients who are older than 55 years have hazard ratios of 1.88 for all patients, 1.93 for male patients and 1.72 for female patients.

Table 1: Model with linear effect estimates of sex and metrical age

Covariate	Hazard Ratio	Std. error	95% Credible Interval Lower	Upper
Male	0.496	0.249	0.296	0.797
Age	1.009	0.006	1.002	1.020

Table 2: Model with linear effect of categorized age for all, male and female patients.

All patients

Covariate	Hazard Ratio	Std. error	95% Credible Interval Lower	Upper
Male	0.465	0.254	0.283	0.774
23-39 years	1.336	0.396	1.624	2.950
40-55 years	1.519	0.375	1.382	3.283
>55 years	1.883	0.358	1.062	3.845

Male patients

Covariate (Age)	Hazard Ratio	Std. Error	95% Credible Interval Lower	Upper
23-39 years	1.557	0.465	1.904	3.333
40-55 years	1.705	0.475	1.382	3.650
>55 years	1.929	0.482	1.539	3.811

Female Patients

Covariate (Age)	Hazard Ratio	Std. error	95% Credible Interval Lower	Upper
23-39 years	1.246	0.413	1.709	3.031
40-55 years	1.626	0.387	1.417	3.789
>55 years	1.724	0.438	1.481	3.600

Table 3: Deviance Information Criterion (DIC) for the various models for combined, male and female patients.

Model	DIC		
	All patients	Male	Female
1	751.534	363.124	416.924
2	749.156	346. 260	408.026
3	742.281	337.341	401.092
4	738.381	329.106	377.028
5	749.704	341.876	389.682
6	742.632	334.405	365.080
7	749.318		
8	721.342	-	-
9	742.354	-	-

In Table 3, values of Deviance Information Criterion (DIC) for the various models under independence and those fitted with frailty are presented. Comparing models with P-splines prior, with linear effect of metrical age and with categorized age, the values of DIC for P-splines models are seen to be generally least, implying best models. As observed, for all patients, when metrical age is fitted with P-splines prior, (i.e Model 3 for independence and Model 4 with frailty), the values of DIC are 742.281 and 737.043 respectively compared with models fitted as linear effect of metrical age which have DIC of 751.534 and 749.146 for independence model (Model1) and model with frailty (Model 2) respectively. The corresponding models under categorized age (Model 5 and Model 6), though performed worse than P-spline models, and are observed to be better in performances (with DIC of 749.704 and 742.632 respectively) than the models with linear effects of metrical age under both independence and frailty specifications. The directions of the performances are the same in gender models. For example, for male patients under independence specification, DIC for P-splines model (Model 3) is 337.341, compared to 363.124 and 341.876 under model for linear effect of metrical age (Model 1) and model for categorized age (Model 5) respectively. Under frailty specification, P-spline model has DIC of 329.106 while models for linear effect of metrical age and categorized age have DIC of 346.260 and 334.405 respectively. Also for female patients, under independence specification, P-splines model has DIC of 401.092,

model with linear effect of metrical age has 416.924 and model with categorized age has 389.682. The corresponding values under frailty model are 377.028, 408.026 and 365.080 for P-splines model, model with linear effect of metrical age and model with categorical age respectively.

Comparing frailty models with independence models (without frailty), it is observed that models fitted with frailty generally perform better than independence models. As observed, for combined, male and female patients, all models fitted with frailty have lower values of DIC than those fitted under independence assumption. These include Model 4 (frailty) versus Model 3 (independence) for P-splines model, Model 2 (frailty) versus Model 1 (independence) for model with linear effect of metrical age and Model 6 (frailty) versus Model 5 (independence) for model with categorized age.

The major interest of modelling interaction between age and sex in this study is to investigate the comparative performances of the various frailty models when interaction terms are added to models fitted with metrical age as linear effect as well as nonlinear effect with P-spline and model fitted with categorical age. Therefore, only the AIC, rather than the hazard ratios are reported in Table 3. As observed, inclusion of multiplicative interaction between sex and metric age, modelled as linear effect (Model 7 with AIC 749.318) does not show any improvement on the corresponding model without interaction term (Model 2 with DIC 749.156). On categorized age, inclusion of interaction term (Model 9, DIC=742.354) only shows slight improvement over the corresponding model without interaction (Model 6, DIC=742.632). For P-spline model, interaction effects of sex with metrical age substantially improves the model, the DIC is 721.342 for model with interaction term (Model 8) and 738.381for model without interaction term (Model 4). However, comparing the models with interaction terms, it is observed that model fitted with P-spline has the best performance while that fitted with metrical age as linear effect has the worst performance.

4. Conclusion

In the analysis of data on hospital admission for the cancer patients under study, results showed significant differences among age groups with respect to the risk of dying from the selected cancer considered. Results of Deviance Information Criterion (DIC) also revealed that when we allowed for non – linearity in the effects on the metrical covariate (age), the model with P-splines prior as in Lang and Brezger (2004) was found to be more adequate than fitting metrical age as linear effect. It was also found that to study the linear effect of age on death from cancer, categorizing age was a better alternative. The study also confirmed that assumption of independence for the survival time of subjects clustered by cancer type was inadequate, rather a model that accounted for heterogeneity induced by such clustering (frailty model in this study) was preferred. On the effect of interaction effect of sex and age on the death due to cancer, model that contained interaction between sex and age when metrical age was modeled nonlinearly was observed to be better than when metrical or categorized age was modeled linearly.

Caveat: The findings in this study are based on the age reported by the patients which may not be their true age. Despite this limitation, the study strength is significant

Acknowledgement

The Author is grateful to the anonymous reviewers for their critical comments and useful suggestions. He is also grateful to the Faculty of Physical Sciences for sponsoring the publication of this Maiden Edition.

References

Abiodun, A.A. (2009). Bayesian Analysis of Right Censored Survival Time Data. Global Journal of Mathematical Sciences. **8**, 13-17.

Adeleke, K.A., Abiodun, A.A. & Ipinyomi, R.A.(2013). On Estimation Methods and Test for Proportional Hazards Assumptions in Survival Data. Scientia Africana. **12**(1), 20-28.

Brezger, A. & Lang, S. (2006). Generalized structured additive regression based on Bayesian P-splines. Comput. Stat. Data Anal. **50**, 967-991.

Breslow, N. E. (1974). Covariance analysis of censored survival data. Biometrika **30** , 89-100.

Callegaro, A. & Iacobelli, S. (2012). The Cox shared frailty model with log-skew-normal Frailtes. Statistical Modeling 2012; **12**(**5**), 399–418.

Cox D.R. (1972). Regression models and Life-tables (with discussion). J. Roy. Statist. Soc. Ser. B. **34**,187-220.

Clayton, D. G. (1991). A Monte Carlo method for Bayesian inference in Frailty Models. Biometrics **47**, 467-485.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. J. Am.

Statist. Assoc. **72**, 557-65.
Eschenbach, A. & Collins, F. (2005). Toward a Comprehensive Genomic Analysis of Cancer. Workshop Executive Summary. The Ritz Carlton Washington, DC.
Fine, J. P., Glidden, D. V., & Lee, K. E. (2003). A simple estimator for a shared frailty regression model. Journal of the Royal Statistical Society Series B **65**, 317–329.
Hennerfeind , A., Brezger, A. & Fahrmeir, L. (2006). Geoadditive Survival Models. Journal of the American Statistical Association, **101**, 1065-1075.
Huang, X.and Liu, L. (2007). A Joint Frailty Model for Survival and Gap Times between Recurrent Events. Biometrics, **63**(2), 389-397.
Jones, G. L. (2004). On the Markov chain central limit theorem. Probability Surveys, **1**, 299-320.
Kim, S & Dey, D.K. (2008). Modeling multilevel survival data using frailty models. Communications in Statistics– Theory and Methods, **37**, 1734–41.
Kom´arek, A. & Lesaffre, E. (2008). Bayesian accelerated failure time model with multivariate doubly-interval-censored data and flexible distributional assumptions. Journal of the American Statistical Association, **103**, 523–33.
Kom´arek, A and Lesaffre, E. (2009). The regression analysis of correlated interval-censored data: Illustration using accelerated failure time models with flexible distributional assumptions. Statistical Modeling, **9**, 299–319.
Lang, S. & Brezger, A. (2004). Bayesian P-splines. Journal of Computational and Graphical Statistics,13: 183-212.
Liu D, Kalbfleisch J.D, and Schaubel, D.E. (2011). A positive stable frailty model for clustered failure time data with covariate-dependent frailty. Biometrics 67 (1): 8-17.
Martinussen T, Pipper CB. (2005). Estimation in the positive stable shared frailty Cox proportional hazards model. *Lifetime Data Analysis* 11(1):99-115.
Ripatti, S. and Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, 56, 1016-1022.
Spiegelhalter, D.J., Best, N.G., Carlin B.P. and Van der Linde, A.(2002). Bayesian Measures of Model Complexity and Fit (with discussion). Journal of Royal Statistical Society.B, *64, 583-640.*
S. Wassertheil-Smoller (2004). *Biostatistics and Epidemiology*. Springer-Verlag New York, Inc.
Xue, X. and Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. Lifetime Data Analysis 2, 277–289.
Yashin ,A.I. and Iachine, I.A. (1995a.). How long can humans live? Lower bound for biological limit of human longevity calculated from Danish twin data using correlated frailty model. Mechanisms of Ageing and Development 80:147-159.
Yau, K.K.W., McGilchrist, C.A. (1997). Use of generalised linear mixed models for the analysis of clustered survival data. Biometrical J. 39, 3–11.