STATISTICS IN TRANSITION new series, Summer 2014
Vol. 15, No. 3, pp. 437–452

JOINT LONGITUDINAL AND SURVIVAL DATA MODELLING: AN APPLICATION IN ANTI-DIABETES DRUG THERAPEUTIC EFFECT

Atanu Bhattacharjee[1], **Dilip C. Nath**[2]

ABSTRACT

The longitudinal and survival analyses are useful tools in the exploration of drug trial data. In both cases the challenge is to deal with correlated repeated observations. Here, the joint modelling for longitudinal and survival data has been carried out via Markov chain Monte Carlo (MCMC) method in type 2 diabetes clinical trials to compare different combinations of drugs, viz. Metformin plus Pioglitazone and Gliclazide plus Pioglitazone. Despite the complexity of the model it has been found relatively easier to implement with WinBugs software. The results have been computed and compared with software R. In both types of the analyses it has been found that no estimates of treatment appear to have significant effect on the evolution of the matter of HBA1c, neither on the longitudinal part nor on the survival one. The Bayesian approach has been considered as an extended tool with classical approach for estimation of clinical trial data analysis.

Key words: random effects, semi-parametric survival model, Weibull distribution, linked sub-models.

1. Introduction

The longitudinal and survival analyses are useful tools in exploring the drug trial data. In type diabetes drug trials, the level of HBA1c is a widely used biomarker for diabetes while studying the efficacy of the drugs in patients. In drug effect comparison the level of HBA1c is used to measure over follow-up periods in clinical trials. The repeated measurements of HBA1c on the same patients give the scope to application of longitudinal and survival data analysis. The level of HBA1c is an important indicator for measuring the endogenous glucose over a period of 2-3 months by recommendation of The International Expert Committee

[1] Lecturer (Biostatistics), Division of Clinical Research and Biostatistics, Department of Cancer Registry and Epidemiology, Malabar Cancer Centre, Kerala-670103, India, E-mail: atanustat@gmail.com.
[2] Professor (Statistics), Department of Statistics, Gauhati University, Guwahati 781014, India.

Report (2009). HBA1c is the important diagnostic parameter for type 2 diabetes by the report of American Diabetes (2010). The mean HBA1c is a powerful predictive tool to determine the diabetes complications are concluded by Lind et al. (2008) and Stratton et al. (2000). The HBA1c is positively associated with blood sugar level has been concluded by DCCT Study Group, 1995. The Bayesian approach in autoregressive longitudinal data analysis in type 2 diabetes patients of India has been explained by Nath and Bhattacharjee (2011). The Bayesian approach has been found the best choice in model variable selection by Nath and Bhattacharjee (2011). The joint model is associated with sub-models by the longitudinal and survival process measurement model concluded by Henderson et al. (2000). In the last two decades, the field of longitudinal and survival data analysis was enriched through adjusting statistical inferences on longitudinal measurements by Carlin et al. (2000), Celeux et al. (2006), Chen (2006), Schluchter (1992), DeGruttola and Tu (1994), Elashoff and Li (2008), Little (1995), Henderson et al. (2000), Hogan and Laird (1997), and many others.

In this context, the linear or random effect model is found more effective by Tsiatis et al. (1995). Li et al. (2009) proposed the joint model for longitudinal and survival data in the correlated repeated observations. Deslandes et al. (2010) concluded that the proportional cause-specific hazard model is the standard regression model of choice to compare the competing risks. However, the Cox analysis is a widely used method for the cause-specific hazard model. In this work, the joint longitudinal and survival models are applied to compare the updated mean value of HBA1c as the effect of different drug treatment.

2. Objective

The aim of this work is to compare the drug treatment effect with the result of HBA1c value during different visits in type 2 diabetes patients. The longitudinal and survival analysis is applied with prior assumption. The performance of a combined drug therapy, i.e., “Metformin with Pioglitazone” and “Gliclazide with Pioglitazone” is compared in reducing the HBA1c level. The Bayesian approach in the separate and joint modelling procedure is applied and compared to drug treatment effect in type 2 diabetes patients.

3. Methods

The linear model presented by Tsitaes et al. (1995) is

$$R_{1i}(g)=Z_{1i}+Z_{2i}(g). \quad (1)$$

The parameter $R_{1i}(g)$ can be obtained by U_{1i} and U_{2i}, where (U_{1i}, U_{2i}) are subject-specific bivariate normal distributions with σ_1^2,σ_2^2 standard deviation. The next term R_{2i} can be segregated to

$$R_{2i}(g)=\lambda_1 Z_{1i}+ \lambda_2 Z_{2i}+ \lambda_3(Z_{1i}+Z_{2i})+Z_{3i}, \text{ where } Z_3\sim N(0,\sigma_3^2) \text{ and} \quad (2)$$

where λ_1 can be taken as a coefficient.

The sequence of the response variables $Y_{i1}, Y_{i2}, \ldots\ldots Y_{in}$ at times $g_{11}, g_{21}, \ldots g_{n1}$ can be obtained from

$$Y_{ij}=\mu_i(g_{ij}) \tag{3}$$

where μ_i is the link function for $g_{ij} \sim N(0,\sigma^2_2)$ which is a sequence of mutually independent measurement errors. It has also been assumed that $\mu_i(g)=x_{1i}(g)'\beta$, in which the vectors $x_{1i}(g)$ and β give the time-varying explanatory variable and their corresponding regression coefficient.

In the case of survival modelling for the time t, the semi-parametric multiplicative model is extended into

$$\tau_i(g)=\tau_0(g)\alpha_0(g)\exp\{x_{2i}(g)'\beta+R_{2i}(g)\}, \tag{4}$$

where $\alpha_0(g)$ is unspecified and X for the covariate information. The term R_{2i} is useful as a latent process. The parameter $\tau_0(g)$ is the baseline hazard function.

3.1. Longitudinal data models

To deal with longitudinal data with continuous outcome the widely used method is the linear mixed effects model. The linear mixed effect longitudinal models have had a long history in biostatistical theory and practice since the first published paper of Laird and Ware (1982). If $Y_{i1}, Y_{i2}, \ldots Y_{ini}$ is i^{th} subject observations for the $g_{i1}, g_{i2}, \ldots, g_{ini}$ times then the model can be formulated to

$$Y_{ij}=\mu_i(g_{ij})+R_{1i}(g_{ij})+\varepsilon_{ij} \tag{5}$$

where $\mu_i(g_{ij}) = x^T_{1i}(g)\beta_1$ is the mean response, $R_{1i}(g_{ij})=d^T_{1i}(g_{ij})Z_i$ is applied to explain the subject-specific random effects, and $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$ is for random error. The terms $R_{1i}(g)$ is applied for subject specific HBA1c observations. The time-varying covariates are explained by the vectors $x_{1i}(g)$ and β_1. The term U_i is used to represent the random factor of the covariates $d_{1i}(s)$ (as compartment of $x_{1i}(g)$) and assumed distributed as $N(0,\Sigma)$.

3.2. Survival data models

The semi-parametric survival model is becoming an attractive tool for the survival analysis. However, the parametric model is more attractive due its simplicity in the survival analysis. The widely applied statistical methods for the survival analysis are Weibull and Cox proportional hazard models.

In the case of the parametric model the i^{th} subject is assumed to follow the Weibull distribution by $g_i \sim$ Weibull $(r, r_i(g))$.

where $$\log(r_i(g)) = x^T_{2i}(g)\beta_2+R_{2i}(g) \text{ and } r > 0. \tag{6}$$

The $x_{2i}(g)$ and β_2 are the covariates of interest and corresponding regression coefficients. The object $R_{2i}(g)$ is applied for the subject specific covariate and intercepts.

However, the event history can be formulated for time g by

$$\tau_i(g)= \tau_0(g)t^{r-1}r_i(g) = \tau_0(g)t^{r-1}\exp(x^T_{2i}(g)\beta_2+R_{2i}(g)), \quad (7)$$

Guo et al. (2004) applied the semi-parametric proportional hazard model in clinical trial by

$$\tau_i(g)= \tau_0(g)\exp(x^T_{2i}(g)\beta_2+R_{2i}(g)), \quad (8)$$

where $\tau_0(g)$ is used for the baseline hazard function. The fundamental properties of the model were discussed by Cox and Oakes (1984).

3.3. Joint model

The joint model has been linked to sub-models by the measurement model for the longitudinal process and the intensity model for the survival process. The connection between longitudinal and survival analysis can be established by stochastic dependence between R_{1i} and R_{2i}. Henderson et al. (2000) discussed the joint modelling via latent zero-mean bivariate. The joint model can be classified into two linked sub-models, (i) the measurement model for the longitudinal process and (ii) the intensity model for the survival process. The joint model becomes applicable to the sub-model.

The joint model in equations (3) and (4) can be formed by

$$R_{1i}(g)=Z_{1i}+Z_{2i}(g), \quad (9)$$

and

$$R_{2i}(g)=\lambda_1 Z_{1i}+\lambda_2 Z_{2i}+\lambda_3(Z_{1i}+Z_{2i}g)+Z_{3i} \quad (10)$$

Equation (3) used the random intercept model as a link function to the longitudinal data.

In equation (9) $(Z_{1i}, Z_{2i})^T$ follows the bivariate normal distribution with $N(0,\Sigma)$, Z_{3i} is independent and assumed to follow $N(0,\sigma^2)$. The parameters λ_1,λ_2 and λ_3 in the survival model (9) measure the association between the two sub-model indicated by the random intercept, slopes and fitted longitudinal value at the even time $R_{1i}(g)$.

The dependence between R_{1i} andR_{2i} is useful to describe the relation between longitudinal and survival processes.

The longitudinal model (3) is basically the random effect model introduced by Laird and Ware (1982). In equation (6), the parameters λ_1, λ_2, and λ_3 are functional to describe the association between two sub-models through random intercepts, via event $R_{1i}(g)$at time t. It is assumed that the latent variables $(Z_{1i},Z_{2i})^T$ have bivariate Normal distribution $N(\mathbf{0},\mathbf{\Sigma})$. More specially, Z_{3i} is assumed with $N(0, \sigma^2_\varepsilon)$. The term U_{3i} is assumed to be not dependent on $(Z_{1i}, Z_{2i})^T$.

4. Analysis of Metformin with Pioglitazone or Gliclazide with Pioglitazone data

4.1. Sources of data

The data set obtained as a secondary source has been taken from the clinical trial conducted in 2008. The patients are taken from the randomized, double blind and a parallel group study conducted in Menakshi Mission Hospital, Tamil Nadu. A total of 65 patients has been selected to participate in the study, 32 in (1) A combination of Metformin with pioglitazone, and 33 in the group of (2) A combination of Pioglitazone with Gliclazide.

4.2. Description of data set

The drug effectiveness is compared through longitudinal and survival data.

Table 1. Description of HBA1c according to different treatment groups and visits

Treatment	Visits	Min	Max	Mean	SD	Missing observation	Available number of observations
Metformin with pioglitazone	HBA1c1st	7.0	12.6	9.52	0.23	0	32
	HBA1c 2nd	6.8	11.7	8.31	0.35	4	28
	HBA1c 3rd	6.3	10.8	7.52	0.29	10	22
Gliclazide with pioglitazone	HBA1c 1st	6.8	12.9	9.51	0.28	0	33
	HBA1c 2nd	6.7	11.9	8.62	0.33	4	29
	HBA1c 3rd	6.3	11	8.03	0.30	13	20

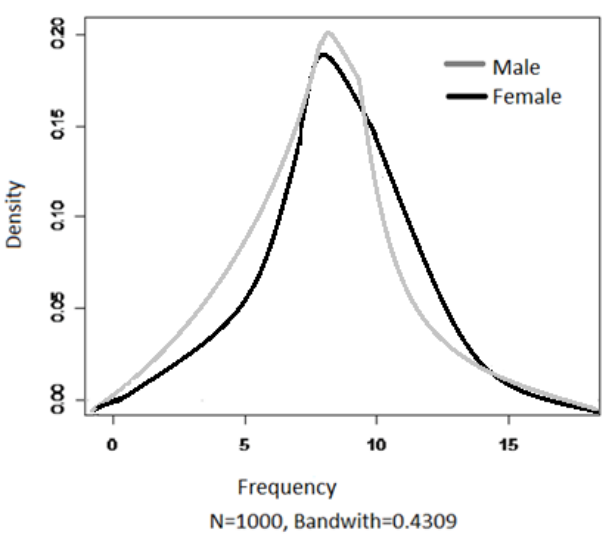

Figure1. Estimated sex wise posterior density of the patient from joint analysis

In this trial, n =65, type 2 diabetes who met the entry conditions were included and randomly allocated to receive either Metformin with Pioglitazone or Gliclazide with Pioglitazone.

The HBA1c levels have been recorded at the study entry, 3and 12 month visits. The death of the subject has also been recorded. However, it is to be noted that the reason of death cannot be specify due to the drug effect. The recorded sample sizes for the drug group (Metformin plus Pioglitazone) in three visits are (32, 28 and 22) and (33, 29 and 20) for the (Pioglitazone plus Gliclazide) group. The estimated posterior density observed adjusted through male and females are given in Figure1. The data is highly affected by drop-out and missing data over time due to the occurrence of death. The Kaplan-Meier curve has been used to show the comparative figure of death between the two drug groups over the follow-up visits. It shows that the survival rate among both groups were same up to the initial 100 days after the randomization. Afterwards, survival in the Pioglitazone with Gliclazide group has been found to be better than Metformin with Pioglitazone group. The level of HBA1c is represented through Y_{ij} for i^{th} observations of the j^{th} individual. The considered dichotomous covariates are Sex (female=0, male=1), value of ECO and ECG (Normal level=0, otherwise 1), and Drug (Pioglitazone with Gliclazide=0 and Metformin with Pioglitazone = 1). The covariates value levelled with "0" is considered as reference value in the analysis.

The objective of the study is to observe the effect of the drug on HBA1c and survival time in type 2 diabetes individuals.

5. Analysis

The analyses for the longitudinal and survival data in type 2 diabetes trial are compared with the Bayesian approach. The linear random effects model for HBA1c is specified as

$$Y_{ij}=\beta_{11}+\beta_{12}*Drug_i+\beta_{14}{}^{*}Sex_i+\beta_{15}{}^{*}ECO_i+\beta_{16}{}^{*}ECGi+R_{1i}(g_{ij})+\varepsilon_{ij} \qquad (11)$$

where $R_{1i}(g_{ij})=Z_{1i}+Z_{2i}g_{ij}$. The term $R_{1i}(g_{ij})$ is induced as a random factor for the intercept and slopes over the duration of study, where the $Z_i=(Z_{1i},Z_{2i})^T \sim N(0,\Sigma)$. It gives the scope to assume that different individuals have different observations before as well during the study of HBA1c.

The estimated regression coefficients have been obtained by R programming. In the case of longitudinal analysis, the rlm (http://cran.r-project.org/web/packages/) function has been applied in R, whereas in the case of survival analysis surv(http://cran.r-project.org/web/packages/) function has been used in survival library. The summarized results are given in the Table 2. As a results the estimated average mean for the Metformin with Pioglitazone is obtained with -0.42 with 95% confidence interval of (-0.67, 0.17), proposing significant increment of HBA1c in the Metformin with Pioglitazone group as compared to Pioglitazone with Gliclazide group.

The comparative changes of HBA1c level are provided in the Figure 3. The estimated regression coefficient value for ECO with 95% confidence interval is observed with 0.23 (-0.01, 0.47). Hence, a patient who has randomized with Drug 1 is found to be more effective to reduce the level of HBA1c in comparison with Drug 0. Other variables ECG and sex are observed with insignificant contribution. Similarly, ECG and ECO are found statistically not significant in survival analysis.

Table 2. Classical analysis for type 2 diabetes drug treatment effect data

Parameters	Point estimate	95% Confidence Interval
Longitudinal Data Analysis (Linear Mixed Effect Model)		
Intercept	9.58	(9.32, 9.06)
SEX(reference=female)	-0.29	(-0.39, -0.19)
ECO(reference=Normal)	0.23	(-0.01,0.47)
ECG(reference= Normal)	0.12	(-0.13,0.37)
DRUG(reference=Pioglitazone with Gliclazide)	-0.42	(-0.67,0.17)
DRUGXTIME	-0.18	(-0.42,0.06)
Survival Analysis		
Intercept	9.55	
SEX	-0.12	(9.12,9.98)
ECO	0.14	(-0.22,-0.03)
ECG	-0.18	(-0.04,0.33)
DRUG	-0.33	(-0.45,0.09)
DRUGXTIME	-0.11	(-0.74,-0.08)
		(-0.24,-0.02)

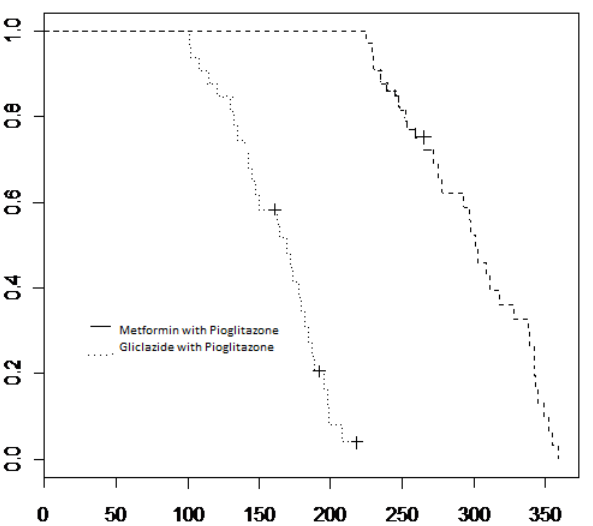

Figure 2. Kaplan-Meier Curve for the drug effect comparison in the type 2 diabetes patients.

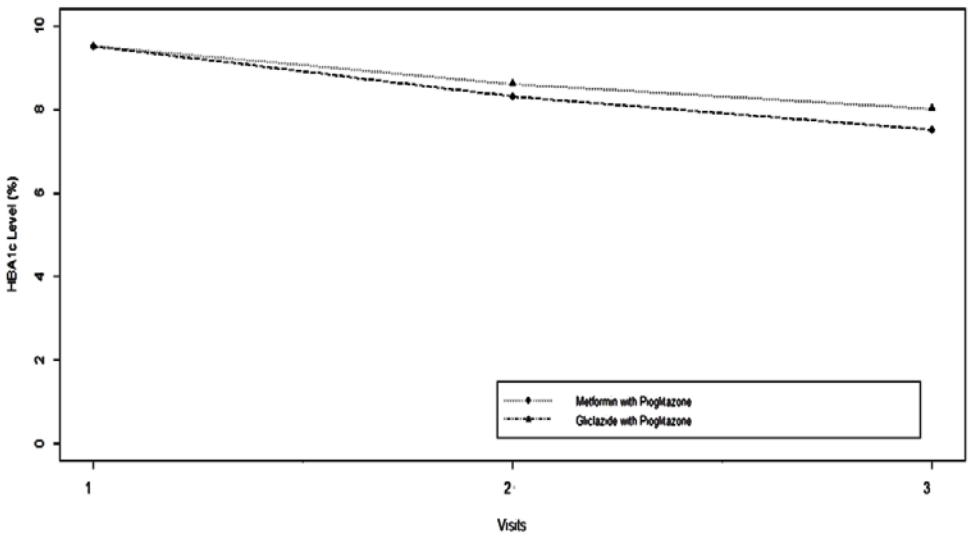

Figure 3. The comparative changes of HBA1c level throughout the study period in drug treatment group.

Henderson et al. (2000) proposed to use the Bayesian approach to fit the joint longitudinal model. The Bayesian approach with the vague prior (Uniform (-1,1)) has been applied and compared with the classical approach. The vague prior has been used to make the possible comparison between the classical approach and the Bayesian approach in WINBUGS. The hyperparameter has been chosen for the minimum impact on the relative data. In the longitudinal sub-model, the multivariate normal and inverse gamma priors have been assumed for the main effect $\boldsymbol{\beta}_1$ and the error variance σ^2_ε, respectively. In the same way the multivariate normal and inverse gamma priors have been assumed for the effects $\boldsymbol{\beta}_2$ and σ^2_ε in the survival sub-model. The vectors $\boldsymbol{\beta}_{1 \text{ and }} \boldsymbol{\beta}_2$ have been expressed by $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{16})^T$ and $\boldsymbol{\beta}_2 = (\beta_{21}, \beta_{22}, \beta_{23}, \beta_{24}, \beta_{25})^T$. The parameters γ_1 and γ_2 have been assumed to follow the normal distribution. Priors are selected to reflect the appearance of likelihood.

6. Model selection

The models under consideration are:

Model 1:- $Y_{ij}=\mu_i(g_{ij})+R_{1i}(g_{ij})+\varepsilon_{ij}$

Model 2- $\tau_i(g)=rt^{r-1}\mu_i(g) = rt^{r-1}\exp(x^T_{2i}(g)\beta_2+R_{2i}(g))$,

Model 3:- $Y_{ij}=\mu_i(g_{ij})+(Z_{1i}+Z_{2i}(g))(g_{ij})+\varepsilon_{ij}$

Model 4:- $\tau_i(g)=rt^{r-1}\mu_i(g) = rt^{r-1}\exp(x^T_{2i}(g)\beta_2+ (\lambda_1U_{1i}+\lambda_2Z_{2i}+\lambda_3(Z_{1i}+Z_{2i}g)+ Z_{3i})$

The comparison between different models is an important issue in the statistical inference. In the case of the Bayesian approach, the widely applied tools are AIC, DIC, BIC and Bayes factor for model comparison. In this work, we have used the DIC. Our priors are selected to make less influence on the likelihood. The model selection can be performed through AIC, BIC, Bayes factor and DIC. Like other selection methods, DIC also gives the model summary to single parameters, through a specific Bayesian inference. Let θ and y be the parameters of interest and the response variable is defined as

$$p_D = E_{\theta/y}[D(\theta)] - D\big(E_{\theta/y}[\theta]\big) = \bar{D} - D(\bar{\theta}) \tag{12}$$

The notation $D(\theta)$ is the deviance function and $D(\theta) = -2\log f(y/\theta) + 2\log g(y)$, where $f(y/\theta)$ is the likelihood function is and $g(y)$ is the standard function of the data. Further, $D(\theta)$ can be formed through $D(\theta) \approx D(\bar{\theta}) + \chi^2_p$. It is formulated through Bayesian Central limit Theorem and details are available in Carlin and Louis (2000). The model selection is obtained through

$$DIC = \bar{D} + p_D \tag{13}$$

Here, p_D is the number of parameters. The posterior expectation of equation (12) is $\bar{D} = E_{\theta/y}[D(\theta)]$, small value of p_D and corresponding minimum value of the DIC gives maximum effective model. The details about the DIC can be seen in the highly cited papers of Spiegelhalter et al. (2002). In WINBUG the parameters are obtained through MCMC technique. There are several versions of DIC available for model selection in recent articles, namely Celeux et al. (2006) and Chen (2006).

The Table 3 gives the DIC values for different models of drug trial comparison data in type 2 diabetes patients. The results are obtained by the two parallel chains of MCMC sampling through 10,000 iterations. We start with simple model of equation (1). As an extension, the term R_{1i} has been added in the equation (1) and in both cases the DIC values have been obtained. The DIC value for the Model (1) is 2345 and for the Model (2) is 2356. In the case of survival analysis, the DIC values for the Model (3) and the Model (4) are found to be 2424 and 2452, respectively. The minimum DIC value of the specific model can be considered as the best fitted model. The details about DIC can be cited with Spiegelhalter et al. (2002). Here, the minimum DIC value for the Model 1 has been found. So, it can be concluded that the Model 1 is the best.

7. Comparison of separate and joint models

The Table 2 gives the point estimates of regression coefficients for covariate of interest by the linear mixed effect model. The linear mixed effect model has

been computed with respect to sex (female=0), Drug (Metformin with Pioglitazone=1). The regression coefficient -0.29 in the case of sex showed that the male type 2 diabetes patients are reduced to lower amount of HBA1c in comparison to the female ones. The Table 3 gives the Highest Posterior Density (HPD) interval for the covariates in different models. The results for the longitudinal model have been obtained from the model in equation (5). The term $R_{1i}(g_{ij})$ is applied as an extension of the joint model to the separate longitudinal model. The term $R_{1i}(g_{ij})$ has also been separated to Z_{1i}, Z_{2i} by equation (9). The 95% credible intervals for the survival model have been obtained from the equation (3) and equation (5) for the survival model, where $R_{2i}(g)$ is used as an extension over the separate model. The performances of both models are found similar. The regression coefficient 0.01 obtained through longitudinal sub-model with separate analysis confirmed that the male type 2 diabetes patients are reduced to higher amount of HBA1c in comparison to the female ones. On the other hand, in the case of sex regression coefficients, the longitudinal sub-model in joint analysis, survival sub-model with separate analysis and joint analysis follow the same pattern as observed in separate analysis in the longitudinal sub-model. The covariates of interest, ECG and time, are observed with considerable extension, while only ECG is found significant in the case of the survival sub-model. The regression coefficients from the classical approach are observed with 0.12(-0.13, 0.37) and 0.23(-0.01,0.47) for ECG and ECO, respectively. In the case of joint modelling applied through the Bayesian approach the posterior means of the regression coefficients are obtained with 0.03(-0.39, 0.42) and -0.02(-0.45, 0.42) for ECG and ECO, respectively.

Table 3. Posterior estimates of the parameters observed through different models

Parameter	Separate analysis			Joint analysis		
	Posterior mean	DIC	95% Credible interval	Posterior mean	DIC	95% Credible interval
Longitudinal Sub-model						
Intercept(β_{11})	8.70	2345	(8.23,9.15)	8.72	2356	(8.21,9.17)
Time(β_{11})	-0.17		(-0.97,0.09)	-0.19		(-0.93,0.07)
Time*Drug(β_{12})	0.19		(-0.14, 0.53)	0.15		(-0.12,0.49)
Sex(β_{13})	0.01		(-0.33,0.37)	0.03		(-0.35,0.39)
ECG (β_{14})	-.02		(-0.41,0.46)	-.03		(-0.39,0.42)
ECO (β_{15})	-0.01		(-0.48,0.45)	-0.02		(-0.45,0.42)
Σ_{11}	1.97		(1.41,2.67)	1.95		(1.38,2.63)
Σ_{22}	0.98		(0.73,1.5)	0.95		(0.70,1.2)
P	-0.12		(-0.30,0.07)	-0.10		(-0.27,0.05)
σ^2	0.95		(0.73,1.21)	0.93		(0.71,1.18)
Survival Sub-model						
Intercept(β_{21})	-15.31	2424	(-19.31,-8.75)	-15.31	2452	(-19.36, -8.76)
Drug(β_{22})	-0.47		(-4.94,4.24)	-0.49		(-4.99,4.28)
Sex(β_{23})	2.81		(-0.75,5.85)	2.80		(-0.79,5.83)
ECG (β_{24})	0.00		(-2.49,2.99)	0.01		(-2.53,3.01)

The separate analysis in longitudinal setup reveals the regression coefficients with -0.02(-0.41, 0.46) and -0.01(-0.48, 0.45) for ECG and ECO. The separated and joint survival analysis is computed with regression coefficients by 0.00(-2.49, 2.99), 0.01(-2.53, 3.01) for ECG and 0.19(-2.72, 2.51), -0.17 (-2.12, 2.15) for ECO, respectively. It is concluded that the regression coefficients obtained through classical approach for ECG are higher in joint and separate approach in longitudinal setup and further followed by survival setup through prior assumption. The same pattern is obtained in the case of ECO. The highest value of the regression coefficient is found with frequency approach. In both types of analysis it is found that no estimates of the treatment appear to have significant effect on the evolution of the matter HBA1c either on the longitudinal part or on the survival. The rate reduction of HBA1c over the follow-up period is found higher in the Metformin with Pioglitazone group.

8. Discussion

In this paper, the Bayesian approach with the longitudinal and survival analysis is applied in the type 2 diabetes drug comparison. This type of the model is important in clinical trial. The models are also useful with other biochemical parameters. It is important to investigate how the biomarker of interest changes over time and its correlation with the treatment under study to better explore the therapeutic effect as pointed by Deslandes and Chevret (2010). The results are obtained through the freely available software and compared with R and WINBUGS. Due to intention-to-treat and other logistical reasons, the whole data set has not been provided to the authors for analysis. The work is carried out only on fully observed but partially data set. Therefore, the whole information about mortality of the patients could not be provided. The aim of this paper is to compare two effects of drug treatment through HBA1c level among type 2 diabetes. Nathan et al. (2009), Holman et al. (2007), Holman et al. (2009) and Meneghini et al. (2007) recommended the level of HBA1c as thresholds for starting insulin. Kilpatrick et al. (2008) discussed broadly the limitation of HBA1c for the screening test. Ginde el al. (2008) and Anand et al. (2003) examined the variation of HBA1c with different demographic characters in the US population. Mirzazadeh et al. (2009) found that the HBA1c can be affected by age distribution. Zahra et al. (2010) concluded that the low HBA1c is a strong evidence to rule out diabetes. However, we acknowledge the deficiency in not including the glucose tolerance test in this work. In addition, as another limitation

in this study, we have not used the life style parameters of the type 2 diabetes patients since some patients cannot be followed or died due to other reason. The analysis becomes complicated due to the presence of dropouts in the data. Thus, the analysis of such type of data by separate analysis may generate biased and inappropriate results whereas the application of joint analysis is useful to deal with dropout observations. Recently, Chi (2006), Williamson et al. (2008) and Li et al. (2009) discussed joint modelling in the longitudinal and survival data analysis. Actually, Guo (2004) has motivated our work to apply the Bayesian approach in longitudinal data analysis to obtain the posterior inference for any parameter. Thus, we have developed a fully Bayesian approach, implemented via MCMC in WINBUGS software. Recently, such a Bayesian approach for joint longitudinal and survival analysis has also been implemented by Li et al. (2009). This work illustrates how the joint model strategy may affect the results. Here, the joint analysis is found inferior in comparison with the separate analysis. It may be due to the presence of other complicated issues in the data set. Lind et al. (2008) concluded that the latent mixed effect is appropriate in the hazard model. In this work, it is found by joint longitudinal and separate analysis that Metformin plus Pioglitazone is equally effective to reduce the HBA1c level as compared to Gliclazide plus Pioglitazone.

9. Conclusions

Here, the HBA1c observations by longitudinal and survival analysis tools are compared with type 2 diabetes patients. The results confirm that the joint modelling approach is a useful tool for longitudinal data analysis, survival analysis and, consequently, for the actual application to the drug effect comparison in clinical trials. The Markov Chain Monte Carlo method is employed to effectively estimate HBA1c values for different visits in type 2 diabetes patients. The applied models can be useful in different fields like oncology, endocrinology and other specific drug research. It is confirmed that the combination of Metformin plus Pioglitazone is equally beneficial to reduce HBA1c level, hence the risk of type 2 diabetes. The Bayesian approach is considered as extending over the Frequency approach on longitudinal and survival data analysis.

Acknowledgements

The authors thank the two anonymous referees for their cautious reading and constructive suggestions which have led to improvement on earlier versions of the manuscript. They are also very grateful to R. Saroj for providing data for analysis and to the International Biometric Conference, Japan 2012, for providing a platform to present the initial version of this work.

REFERENCES

AMERICAN DIABETES ASSOCIATION (2010).Diagnosis and classification of diabetes mellitus. Diabetes Care, 33 (Suppl 1), S62–S69.

ANAND, S. S., RAZAK, F., VUKSAN, V., GERSTEIN, H. C., MALMBERG, K., YI., Q., TEO, K. K., YUSUF, S., (2003). Diagnostic strategies to detect glucose intolerance in a multiethnic population, Diabetes Care, 26(2), 290–296.

CARLIN, B. P., LOUIS, T. A., (2000). Bayes and Empirical Bayes Methods for Data Analysis (2nd ed.), Boca Raton, FL: Chapman and Hall/CRC Press.

CELEUX, G., FORBES, F., ROBERT, C. P., TITTERINGTON, D. M., (2006). Deviance information criteria for missing data models, Bayesian Analysis, 4, 651–674.

CHEN, M. H., (2006). Comments on article by celeux et al., Bayesian Analysis, 4, 677–680.

CHI, Y. Y., IBRAHIM, J. G., (2006). Joint models for multivariate longitudinal and multivariate survival data. Biometrics. 62(2), 432–445.

COX, D. R., D. OAKES, (1984). Analysis of Survival Data, London: Chapman and Hall.

DCCT STUDY GROUP, (1995). The relationship of glycemic exposure (HbA1c) to the risk of development and progression of retinopathy in the diabetes control and complications trial, Diabetes, 44(8), 968–983.

DEGRUTTOLA, V., TU, X. M., (1994). Modeling progression of cd4 lymphocyte count and its relationship to survival time, Biometrics, 50, 1003–1014.

DESLANDES, E., CHEVRET, S., (2010). Joint modeling of multivariate longitudinal data and the dropout process in a competing risk setting: Application to ICU data, BMC Med Res Methodology, 2010, 10, 69.

ELASHOFF, R., LI, G., LI, N., (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types, Biometrics, 64, 762–771.

GHAZANFARI, Z., HAGHDOOST, A. K., ALIZADEH, S. M., ATAPOUR, J., ZOLALA, F., (2010). Comparison of HbA1c and Fasting Blood Sugar Tests in General Population. International Journal of Preventive Medicine, 1(3), 187–194.

GINDE, A. A., CAGLIERO, E., NATHAN D. M., CAMARGO, C. A., J. R., (2008). Value of risk stratification to increase the predictive validity of HbA1c in screening for undiagnosed diabetes in the US population. J Gen Intern Med, 23(9), 1346–1353.

GUO, X. U., CARIN BRADLEY, P., (2004). Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages, The American Statistician, February 2004, Vol. 58, No. 1, 1–9.

HENDERSON, R., DIGGLE, P. J., DOBSON, A., (2000). Joint Modeling of Longitudinal Measurements and Event Time Data, Biostatistics, 1, 465–480.

HOGAN, J. W., LAIRD, N. M., (1997). Model-based approaches to analysing incomplete longitudinal and failure time data, Statistics in Medicine, 16, 259–272.

HOLMAN, R. R., THORNE, K. I., FARMER, A. J., DAVIES, M. J., KEENAN, J. F., PAUL, S., LEVY, J. C., (2007). Addition of biphasic, prandial, or basal insulin to oral therapy in type 2 diabetes. N Engl J Med, 357, 1716–1730.

HOLMAN, R. R., FARMER, A. J., DAVIES, M. J., LEVY, J. C., DARBYSHIRE, J. L., KEENAN, J. F., PAUL, S. K., (2009). Three-year efficacy of complex insulin regimens in type 2 diabetes. N Engl J Med, 361, 1736–1747.

KILPATRICK, E. S., (2008). Haemoglobin A1c in the diagnosis and monitoring of diabetes mellitus, J Clin Pathol, 61(9), 977–82.

LITTLE, R. J. A., (1995). “Modeling the drop out mechanism in repeated measures studies,” Journal of the American Statistical Association, 90, 1112–1121.

LIND, M, ODEN, A., FAHLÉN, M., ELIASSON, B., (2008). A Systematic Review of HbA1c Variables Used in the Study of Diabetic Complications, Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 282–293.

LAIRD, N. M., WARE, J. H., (1982). Random-Effects Models for Longitudinal Data, Biometrics, 38, 963–974.

LI, L., HU, B., GREENE, T., (2009). A semi-parametric joint model for longitudinal and survival data with application to hemodialysis study. Biometrics. 65(3), 972–986.

MENEGHINI, L. F., ROSENBERG, K. H., KOENEN, C., MERILAINEN, M. J., LÜDDEKE H. J., (2007). Insulin detemir improves glycaemic control with less hypoglycaemia and no weight gain in patients with type 2 diabetes who were insulin naive or treated with NPH or insulin glargine: clinical practice experience from a German subgroup of the predictive study. Diabetes Obes Metab. 9, 418–427.

MIRZAZADEH, A., BARADARAN, H. R., HAGHDOOST, A. A., SALARI, P., (2009). Related factors to disparity of diabetes care in Iran, Med Science Monit, 15(5), H32–H36.

NATHAN, D. M., BUSE, J. B., DAVIDSON, M. B., FERRANNINI, E., HOLMAN, R. R., SHERWIN, R., ZINMAN, B.,(2009). Medical management of hyperglycemia in type 2 diabetes: a consensus algorithm for the initiation and adjustment of therapy: a consensus statement of the American Diabetes Association and the European Association for the Study of Diabetes, Diabetes Care. 32, 193–203.

NATH, D. C., BHATTACHARJEE, A., (2011). A Bayesian Approach in Autoregressive models in Longitudinal Data Analysis: An Application to Type 2 Diabetes Drug Comparison, Asian Journal of Applied Science, 4(6), 640–648.

SCHLUCHTER, M. D., (1992). Methods for the analysis of informatively censored longitudinal data, Statistics in Medicine, 11, 1861–1870.

SPIEGELHALTER, D. J., NICOLA, G. BEST, CARLIN, B. P., LINDE, A. V. D., (2002). Bayesian measures of model complexity and fit .Journal of the Royal Statistical Society: Series B (Statistical Methodology. 64, 4, 583–639.

STRATTON, I. M., ADLER, A. I., NEIL, H. A., MATTHEWS, D. R., MANLEY, S. E., CULL, C. A., HADDEN, D., TURNER R. C., HOLMAN, R. R., (2000). Association of glycaemia with macrovascular and

microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study, British Medical Journal, 2000; Aug 12; 321(7258), 405–412.

The International Expert Committee (2009).The International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes. Diabetes Care, 32(7), 1327–1334.

TSIATIS, A. A., DEGRUTTOLA, V., ANDWULFSOHN, M. S., (1995). Modeling the Relationship of Survival to Longitudinal Data Measured with Error, Applications to Survival and CD4 Counts in Patients with AIDS, Journal of the American Statistical Association, 90, 27–37.

WILLIAMSON, P. R., Kolamunnage-Dona, R., Philipson, P., Marson, A. G., (2008). Joint modelling of longitudinal and competing risks data. Statistics in Medicine. 27(30), 6426–6438.