Bayesian Joint Modelling of Disease Progression and Time to Death Event of HIV/AIDS Patients under ART Follow up at Shashemene Referral Hospital, Ethiopia

Abstract

Longitudinal data and survival data frequently arise together in practice. Joint modelling refers to the statistical analysis of the longitudinal and survival data while taking account of any association between the repeated measurement and time to event outcomes. The population of our study is all HIV/AIDS patients who were under the follow up of ART at Shashemene referral hospital from January, 2006 to December, 2012. The purpose of this study is to develop separate and joint statistical models in Bayesian framework for longitudinal and survival data of HIV/AIDS patients under ART follow up at Shashemene referral hospital. The WinBUGS software was used to produce MCMC samples from the posterior distribution of the parameters of a desired model. The results of both the separate and joint analyses are consistent. The overall performance of separate model in terms of model goodness of fit is better than joint model. The joint model $M9$ is somehow the simpler (less complex) model as compared to the separate model for this study. Both in joint and separate model linear and quadratic time, sex and tobacco addiction are statistically significant at 0.05 level of significance. Out of the covariates included in the survival submodel of the joint model, knowledge of ART and condom use are found to be significantly associated with time to death at 0.05 level of significance. Also, the results show that the patients’ survival in the ART treatment is not significantly associated with patient specific CD4 fluctuations.

Keywords: Joint model, Longitudinal data, Survival data, CD4 count, ART.

1 Introduction

Many scientific investigations generate longitudinal and survival data, as well as additional covariate information at the same time. A typical example is that of HIV clinical trials, in which a biomarker such as CD4 lymphocyte
count is measured intermittently and time to progression to AIDS or death is also recorded, with possible early dropout or failure to experience event by the end of study (Yao, 2007). Data analysis can mainly focus on either the longitudinal data or the survival data or both. When the analysis focuses on longitudinal data, we often need to address informative dropouts since dropouts are very common in longitudinal studies. When the analysis focuses on survival data, we often need to incorporate time-dependent covariates such as CD4 since the times to event may be associated with the covariate trajectories. Sometimes, the main interest may lie in the association between the longitudinal and survival data. In any of these cases, joint models are required to feature correlated longitudinal and survival data (Wu et al., 2012).

The development of joint model has greatly expanded the scope of models to accommodate many data complexities, yet relatively little attention has been paid to these approaches properties and performance. For our case, we applied this technique to data on HIV/AIDS patients to investigate the patterns of CD4 changes and to characterize the relationship between CD4 features and time to death event. The longitudinal process of CD4 together with time to death has been investigated using models with focus on the association between the longitudinal and survival data.

The central research questions are: What are the factors for determining the longitudinal evolution of CD4 cell count of HIV/AIDS patient under ART follow up? What are the risk factors for the death of HIV/AIDS patient under ART? What is the effect of the longitudinal evolution of CD4 cell count in the time to death of a patient under ART treatment? How strong is the association between the CD4 cell count and the time to death of a HIV/AIDS patient under ART? The objective of the study is to analyze and model CD4 progression and time to death of HIV/AIDS patients based on data at Shashemene referral hospital. The results of such study is very useful in the development of an effective antiretroviral therapy (ART) and monitoring system.

2 Methods

The population of our study is all HIV/AIDS patients who were 16 years old or older and under ART follow up between January, 2006 and December, 2012 at Shashemene Referral Hospital. All patients who were below 16 years and those patients who started ART before January 2006 or after December, 2012 were not included in the study. Patients who were observed for less than three times are also excluded. So our study population consists of 2413 patients who fulfilled the inclusion criteria.

Response variables: Two outcome variables are considered in this study. The first response variable is longitudinal CD4 count. The longitudinal continuous variable is the square root of the CD4 cell counts per mm$^3$ of blood. It is measured repeatedly for each HIV/AIDS patient under ART. The other response variable is the survival outcome variable. It is time to death event of the patient under ART follow up.
**Explanatory variables:** several variables are expected to be predictors of the response variables. For this study, five variables (condom use, number of living room, knowledge, TB status and number of opportunistic infections) are selected to be included in the survival model while observation time, sex of patient, tobacco addiction, functional level, alcohol addiction and number of opportunistic infections are selected to be included in longitudinal model (linear mixed effects model).

### 2.1 Longitudinal Data Analysis

Longitudinal data are measurements of a response variable taken on the same individuals over several occasions. A longitudinal study is defined as a study in which the response for each experimental unit in the study is observed on two or more occasions. The defining feature of a longitudinal data set is repeated observations on experimental units. Longitudinal data require special statistical methods because the set of observations on one subject tends to be intercorrelated (Viviani, 2012). Here CD4 counts have been measured over time for each individual patients (AIDS patients). It is denoted by \( Y_{ij} \) where \( i \) represent the patients and \( j \) represents observation time.

For longitudinal data, there are two sources of variations: within-subject variation (the variation in the measurements within each subject) and between-subject variation (the variation in the data between different subjects). Modeling within-subject variation allows studying changes over time, while modeling between subjects variation allows understanding differences between subjects.

**Linear Mixed Effects Model (LMEM)**

Linear Mixed effects models (LMEM) are widely used in which random effects are introduced to incorporate the between subjects variation and within subject correlation in the data. Model for longitudinal data is:

\[
y_{ij} = \mu_i(s_{ij}) + W_{1i}(s_{ij}) + \epsilon_{ij}
\]

Where \( \epsilon_{ij} \sim N(0, V_i) \), \( \log(V_i) \sim N(\mu_v, \sigma_v^2) \). This model incorporates subject-specific variances, \( V_i \) which is random by itself. \( \epsilon_{ij} \) can not have homogeneous variance. Thus, here, \( V_i \) represents the (true) within-subject variability which follows a log-normal distribution with mean \( \mu_v \) and variance \( \sigma_v^2 \).

### 2.2 Survival Data Analysis

Survival Analysis typically focuses on time to event data. Survival time random variables are always non-negative, i.e if we denote the survival time by \( T \), then \( T \geq 0 \). \( T \) can either be discrete (taking a finite set of values) or continuous (defined on \((0, \infty)\)). Here we use Weibull based models
Parametric Survival Models

Parametric models are models requiring the specification of a probability distribution for the survival times, i.e.,
parametric models assume that the survival data follow some probability distribution. In such models, the base-
line hazard function, \( \lambda_0(t) \), is modeled parametrically. The most commonly used parametric model is the Weibull
model. In the Weibull model, we assume that the survival time for the \( i^{th} \) subject follows a Weibull distribution,
\( t_i \sim \text{Weibull}(p, \mu_i(t)) \).

\[
\log(\mu_i(t)) = X_2^T i(t) \beta_2 + W_2 i(t), \quad p > 0
\]

where the vectors \( X_2(t) \) and \( \beta_2 \) represent (possibly time-dependent) explanatory variables and their correspond-
ing regression coefficients. They may or may not have elements in common with \( X_1(t) \) and \( \beta_1 \) in the longitudinal
model. The form of \( W_2(t) \) is similar to \( W_1(s) \), including subject-specific covariate effects and an intercept (often
called a frailty). The event intensity (or hazard) at time \( t \) is given as

\[
\lambda_i(t) = pt^{p-1} \mu_i(t) = pt^{p-1} \exp\{X_2^T i(t) \beta_2 + W_2 i(t)\}
\]

which is monotone in \( t \) (decreasing if \( p < 1 \), increasing if \( p > 1 \)) and reduces to the exponential (constant in \( t \))
hazard if \( p = 1 \).

2.3 Joint Model

This study used the joint modeling approach developed by Guo and Carlin (2004). Association between the longi-
tudinal and survival processes can arise in two ways. One is through common explanatory variables and the other
is through stochastic dependence between \( W_2(t) \) and \( W_1(s) \).

The joint model consists of two linked submodels, the longitudinal process measurements model and the survival
process model. We can apply this joint modeling strategy to connect the classical models for longitudinal data and
survival data with each other. When association between the two processes exists, we should obtain less biased
and more efficient inferences by using this joint model. Specifically, our joint model links longitudinal (1) and
survival (2) models by taking

\[
W_1 i(s) = U_1 i + U_2 i \ast s
\]

and

\[
W_2 i(t) = \gamma_1 U_1 i + \gamma_2 U_2 i + \gamma_3 (U_1 i + U_2 i) + U_3 i
\]

Where form of the association function, \( W_2(t) \), is similar to \( W_1(s) \), including subject specific covariate effects and
an intercept (often called a frailty). We adopt the usual joint modeling assumption that the \( W_2(t) \) induce all of
the association between longitudinal process and survival time. When \( W_2(t) = 0 \), there is no association between
longitudinal and survival processes. More specifically, we used the joint model that links the linear mixed effects model that incorporates subject-specific variance and the Weibull survival model.

The longitudinal submodel is the LMEM that includes subject-specific heterogeneous variance with each patient receiving random intercept and linear slope terms. The form in $W_1(s)$ is linear in $s$, which is motivated while exploring the longitudinal data. The parameters, $\gamma_1$, $\gamma_2$ and $\gamma_3$ in the survival submodel measure the association between the two submodels induced by the random intercepts and linear slopes and fitted longitudinal value at the event time $W_{1i}(t)$, respectively. As mentioned before, the pair latent variable $(U_{1i}, U_{2i})$ has a mean-zero bivariate Gaussian distribution $\mathcal{N}(0,\Psi)$, while the $U_{3i}$ are independent frailty terms, modeled as iid $\mathcal{N}(0,\sigma^2_3)$, independent of $(U_{1i}, U_{2i})$ and the subject-specific variances $V_i$ have a lognormal distribution $\log(V_i) \sim (\mu_v,\sigma^2_v)$. Regarding the association function, $W_{2i}(t)$, a variety of several latent processes are considered. Finally, the precise nature of the two submodels i.e., the exact form of $W_{1i}(s)$ and $W_{2i}(t)$ and their latent association are selected using Deviance Information Criteria (DIC).

### 2.4 Bayesian Method

The standard maximum likelihood method involves integrating out latent variables from the log likelihood function which is difficult when dealing with high-dimensional variables (Xin et al., 2009). As a result, the proposed joint models are estimated under a Bayesian framework using Markov chain Monte Carlo (MCMC) methods with Gibbs sampling using the non-commercial software WinBUGS. Bayesian joint models have also been studied by various authors, including Faucett and Thomas (1996), Xu and Zeger (2001), Wang and Taylor (2001), Law et al. (2002), Ibrahim et al. (2004), and Huang et al. (2011). Joint models may contain many unknown parameters, which may lead to potential problems in inference.

A main advantage of Bayesian methods is that they can incorporate additional information from similar studies or from experts guess to the model in the forms of prior distributions. Thus, Bayesian methods can be very useful for inference of joint models. For Bayesian joint models, the model parameters are assumed to follow some prior distributions, and inference is then based on the posterior distribution given the observed data. Making use of the usual joint modeling assumption that the subject-specific latent variable induce all of the association between longitudinal process $Y$ and survival outcome $T$, so that $Y$ and $T$ are conditionally independent given random effects $U_i$. Given the random effects, the longitudinal process is assumed to be independent from the event times. So that the full joint distribution of the longitudinal continuous response and time to event can be specified in the form of:

$$f(Y, T, \delta | \theta_1, \theta_2) = \int f(Y|\theta_1, U_i) f(T, \delta | Y, \theta_2, U_i) f(U_i) dU_i$$

(6)
The Likelihood Function

The likelihood function of the above full joint distribution of the longitudinal continuous response and time to event of interest is given as.

\[
L(Y, T, \delta | \theta_1, \theta_2) = \prod_{i=1}^{n} \int f(Y | \theta_1, U_i) f(T, \delta | Y, \theta_2, U_i)^{\delta_i} X(1 - F(T, \delta | Y, \theta_2, U_i))^{1-\delta_i} f(U_i) dU_i
\]

(7)

where

- \(U_i = \{U_{1i}, U_{2i}, U_{3i}\}\) represents the shared underlying process,
- \(\theta_1 = \{\beta_1, \Psi, \mu, \sigma_2^2\}\) are the population parameters as given in the linear mixed effect model.
- \(\theta_2 = \{\beta_2, \gamma_1, \gamma_2, \gamma_3, \sigma_3^2\}\) are the population parameters as given in survival models
- \(f(.)\) and \(F(.)\) denote density and distribution functions, respectively.

Prior Distribution

In a Bayesian approach, model parameters are treated as random variables and assigns probability to each, which is the major difference to the likelihood approach. The assumed distributions for the parameters are called prior distributions. In the study, we have different parameters \(\beta_1\) and \(\beta_2\) the vector of coefficients of covariates in the longitudinal submodel and the vector of coefficients of covariates in the survival submodel, respectively. We used normal distribution as prior distribution for each of parameters in the two vectors. The shape parameter (p) of Weibull model and the association parameters \(\gamma_1, \gamma_2\) and \(\gamma_3\) are assumed to follow gamma distributions.

Posterior Distribution

Bayesian estimation and inference is based on the posterior distribution which is the conditional distribution of unobserved quantities given the observed data. The joint posterior distribution for all unknown parameters \(\theta\) and random effects \(U\) is then given by

\[
f(\theta, U | Y, T) = \frac{f(Y, T | \theta, U) \pi(\theta, U)}{\int f(Y, T | \theta, U) \pi(\theta, U) d(\theta, U)}
\]

(8)

where

- \(f(\theta, U | Y, T)\) is the posterior probability distribution.
- \(f(Y, T | \theta, U)\) is the likelihood function and
- \(\pi(\theta, U)\) is the prior probability distribution.
In the Bayesian framework, inference follows from the full posterior distribution. Bayesian joint model inference is then based on samples drawn from the posterior distribution using an MCMC algorithm such as the Gibbs sampler and Metropolis Hastings. For example, the posterior means and variances of the parameters can be estimated based on these samples, and Bayesian inference can then be based on these estimated posterior means and variances. This sampling can be done using winBugs software.

The WinBUGS software package is a program that performs the Markov chain Monte Carlo algorithms on models and data that are input by the user. The WinBUGS software produces MCMC samples from the posterior distribution of the parameters of a desired model using Gibbs sampler and Metropolis Hastings algorithms. In order to implement MCMC, it is necessary to derive the distribution of each of the model parameters, conditional on all the other parameters. One of the advantages of the WinBUGS software package is that the full conditionals are computed automatically once the user specifies the model. If a parameter distribution is proportional to some known, standard distribution, then it must in fact have that distribution. This is because a probability distribution must always integrate to 1. If a distribution is not standard, then we use the Metropolis-Hastings algorithm, which always involves evaluating the full conditionals in both the numerator and denominator of the acceptance probability, meaning the constant drops out. Hence, in either case, we need only specify the full conditionals up to a multiplicative constant.

3 Results

Two response variables were considered. One response variable is longitudinal response variable. The longitudinal response is the number of CD4 counts per $mm^3$ of blood which were measured approximately every 6 months. Hence, the continuous longitudinal response variable is the square root of number of CD4 counts. The average number of baseline CD4 count is 156.58 per $mm^3$ of blood with standard deviation of 92.535. The longitudinal CD4 measurement distribution is right skewed. But, after transformation by taking square root of the measurement, the distribution of square root of CD4 count be come normal.

The survival endpoint of interest is death of patient. Those patients who missed contact for all study time interval and who were under the follow up till the end of study time were considered as censored. The patients transferred to other hospitals were excluded from study while the transfer in patients were included in the study. From 354 patients included in the study, 333(94.1%) are censored while 21(5.9%) are dead. The time-to-death in months was calculated by subtracting the date of ART start from date of death or censoring.
3.1 Results of Linear Mixed Effects Model Analysis

In any data analysis, before directly going to the analysis first assumptions of the model must be checked. Let $y_{ij}$ denote the square root of $i^{th}$ patient CD4 count value ($i = 1, 2, ..., 354$) at the $j^{th}$ observation time ($j = 1, 2, ..., n_i \leq 11$). $y_{ij} \sim N(\mu_y, \tau_y)$ for conventional LME model and $y_{ij} \sim N(\mu_y, \tau_{V_i})$ and $\log(V_i) \sim N(\mu_v, \tau_v)$ for LME model with heterogeneous within-subject variance for the CD4 counts. Hence, the linear random effects model for square root of CD4 counts is specified as:

$$
\mu(y_{ij}) = \beta_{11} + \beta_{12} obt_{ij} + \beta_{13} obt_{ij}^2 + \beta_{14} sex_i + \beta_{15} fun_i + \beta_{16} alcol_i + \beta_{17} tobac_i + \beta_{18} oi s_i + W_{1i}(obt_{ij})
$$

(9)

where $W_{1i}(obt_{ij}) = U_{1i} + U_{2i} obst(ij)$. Here, $W_{1i}(obt_{ij})$ includes the random effects for intercept and linear time slopes over time. where, $U_i = (U_{1i}, U_{2i})^T \sim N_2(0, \Sigma)$. This specification allows different subjects to have different baseline CD4 counts and different time trends for CD4 counts during the trial.

After determining the mean response of the longitudinal model to be quadratic in time, then in order to examine whether the assumption of heterogeneous within-subject variance for the CD4 counts is supported and also identify the random effects (random intercepts and random linear time slopes) to be included in the model, longitudinal model is fitted using WinBUGS. Table 1 below present the posterior means, standard deviations, Monte Carlo errors and 95% credible intervals for the population parameters of the LME model with heterogeneous within-subject variance for the CD4 counts.

The results show that among the covariates included in the model, linear and quadratic time slopes, gender and tobacco addiction are statistically significant at 5% level of significance and functional level, alcohol addiction and number of opportunistic infections (OIs) are statistically insignificant at 5% level of significance. In Bayesian sense, the 95% posterior credible intervals for both the linear and quadratic time effects and gender excludes 0 while that of functional level, alcohol addiction and number of opportunistic infections (OIs) inclde 0. These estimates, $\beta_{12}$ and $\beta_{13}$, show that on average the longitudinal CD4 measurement significantly increases over time, but the trend is not linear. It displays a fast increase initially and then slows down gradually as time goes by.

In the table, the estimated mean subject-specific precision is $1/\hat{\sigma}_v^2 = 2.158$ with 95% credible interval (1.596, 2.897). Hence, it supports the assumption of heterogeneous variance for the repeated CD4 measurements. Hence, we use the linear mixed effect model that incorporate subject-specific variances for our joint model estimation is acceptable.
Table 1: Results of LME Model that incorporates Patient-Specific Variances Analysis

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Posterior Mean</th>
<th>st.dev</th>
<th>MC error</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\beta_1$)</td>
<td>13.670</td>
<td>0.3171</td>
<td>0.0043</td>
<td>(13.05, 14.29)</td>
</tr>
<tr>
<td>Obstime ($\beta_2$)</td>
<td>2.149</td>
<td>0.0834</td>
<td>0.0007</td>
<td>(1.987, 2.314)</td>
</tr>
<tr>
<td>(Obstime) ($\beta_3$)</td>
<td>-0.118</td>
<td>0.0010</td>
<td>0.0000</td>
<td>(-0.137, -0.099)</td>
</tr>
<tr>
<td>Sex ($\beta_4$)</td>
<td>-0.915</td>
<td>0.3862</td>
<td>0.0059</td>
<td>(-1.677, -0.160)</td>
</tr>
<tr>
<td>Functional ($\beta_5$)</td>
<td>-0.422</td>
<td>0.3331</td>
<td>0.0047</td>
<td>(-1.073, 0.231)</td>
</tr>
<tr>
<td>Alcohol ($\beta_6$)</td>
<td>0.932</td>
<td>0.5627</td>
<td>0.0092</td>
<td>(-2.023, 0.180)</td>
</tr>
<tr>
<td>Tobacco ($\beta_7$)</td>
<td>1.119</td>
<td>0.5610</td>
<td>0.0092</td>
<td>(0.026, 2.212)</td>
</tr>
<tr>
<td>OIS ($\beta_8$)</td>
<td>0.050</td>
<td>0.0890</td>
<td>0.0014</td>
<td>(-0.124, 0.225)</td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tau1</td>
<td>0.106</td>
<td>0.1061</td>
<td>0.0000</td>
<td>(0.087, 0.128)</td>
</tr>
<tr>
<td>tau2</td>
<td>2.381</td>
<td>0.3410</td>
<td>0.0035</td>
<td>(1.793, 3.123)</td>
</tr>
<tr>
<td>$\mu_v$</td>
<td>2.033</td>
<td>0.0532</td>
<td>0.0005</td>
<td>(1.927, 2.136)</td>
</tr>
<tr>
<td>$\tau_v$</td>
<td>2.158</td>
<td>0.3345</td>
<td>0.0048</td>
<td>(1.596, 2.897)</td>
</tr>
<tr>
<td>DIC</td>
<td>12359.700</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2 Results of Weibull Model Analysis

We used backward LR an automatic variable selection method in SPSS to determine the variables to be included in the survival models. Accordingly, of all variables considered, five variables (number of living room, TB status, condom use, knowledge of ART and number of opportunistic infections) are selected to be included in the survival model. After determining the survival time distribution to be Weibull, the survival data is then analyzed using WinBUGS. In a Weibull model, we assume that the survival time for the $i^{th}$ subject follows a Weibull distribution, $t_i \sim Weibull(p, \mu(t_i))$, $p > 0$. The full Weibull Survival Model used in WinBUGS is:

$$\log(\mu(t_i)) = \beta_{21} + \beta_{22}tb_i + \beta_{23}know_i + \beta_{24}cond_i + \beta_{25}ois_i + \beta_{26}rom_i$$  \hspace{1cm} (10)

This is the parameterization used in WinBUGS. From Table 2 below, it is easy to observe the parameter estimates of the full Weibull model. In this model, among the five covariates included in the model, three variables; TB status, knowledge of ART and condom use are statistically significant at 0.05 level of significance. In Bayesian sense, the 95% posterior credible intervals for coefficients of TB status, knowledge of ART and condom use excludes 0 while that of number of living room and number of opportunistic infections (OIs) include 0.

Table 2: Results Weibull Model Analysis

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Posterior Mean</th>
<th>st.dev</th>
<th>MC error</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept ($\beta_{21}$)</td>
<td>-16.360</td>
<td>0.6982</td>
<td>0.0276</td>
<td>(-17.750, -15.040)</td>
</tr>
<tr>
<td>TB status ($\beta_{22}$)</td>
<td>1.182</td>
<td>0.1431</td>
<td>0.0020</td>
<td>(0.899, 1.459)</td>
</tr>
<tr>
<td>Knowledge ($\beta_{23}$)</td>
<td>-0.708</td>
<td>0.0704</td>
<td>0.0009</td>
<td>(-0.845, -0.570)</td>
</tr>
<tr>
<td>Condomuse ($\beta_{24}$)</td>
<td>1.278</td>
<td>0.1322</td>
<td>0.0022</td>
<td>(1.020, 1.540)</td>
</tr>
<tr>
<td>OIS ($\beta_{25}$)</td>
<td>-0.023</td>
<td>0.0288</td>
<td>0.0002</td>
<td>(-0.079, 0.030)</td>
</tr>
<tr>
<td>Living room ($\beta_{26}$)</td>
<td>0.053</td>
<td>0.0473</td>
<td>0.0004</td>
<td>(-0.146, 0.037)</td>
</tr>
<tr>
<td>$p$ (Shape parameter)</td>
<td>3.979</td>
<td>0.1625</td>
<td>0.0064</td>
<td>(3.671, 4.305)</td>
</tr>
</tbody>
</table>
3.3 Model Selection for the Joint Models

The literature on model selection for joint models is quite limited. In practice, the best longitudinal model can be selected based on the observed longitudinal data, and the best survival model can be selected based on the survival data, using standard model selection procedures for these models. Then, we specify reasonable link between the two models, such as shared random effects. As mentioned above, we have chosen the precise nature of the two sub models; the longitudinal to be LME model with subject-specific variances and the survival model to be Weibull. Hence, their association is selected via the DIC Deviance Information Criterion and a hierarchical modeling generalization of the AIC (Akaike Information Criterion). Thinking of $\theta$ and $Y$ as the entire collections of model parameters and data respectively,

$$DIC = \bar{D} + pD$$

(11)

Here, the fit of a model is summarized in the first term by the posterior expectation of the deviance, $\bar{D} = E_{\theta|Y}[D]$ while the complexity of the model is captured in the second term by the effective number of parameters $pD$ Spiegelhalter et al (2002) show $pD$ turns out to be reasonably defined

$$pD = E_{\theta|Y}[D] - D(E_{\theta|Y}[^{\bar{\theta}}]) = \bar{D} - D([\bar{\theta}])$$

(12)

Since small values of $\bar{D}$ indicate good fit while small values of $pD$ indicate a parsimonious model small values of the sum ($DIC$) indicate preferred models. Several joint models with different form of latent processes are explored in order to identify the joint model that fit data well. In all cases, the results are based on three parallel MCMC sampling chains of 50,000 iterations each, following a 25,000 iteration burn-in period. By default, WinBUGS provides the components of DIC for the two submodels (i.e., the terms in the log-likelihood arising from longitudinal and survival model components) to evaluate their relative contributions to the total DIC score; hence the DIC for the longitudinal and survival sub models are denoted as $DIC_1$ and $DIC_2$, respectively.

The table 3 reports $\bar{D}$, $pD$ and $DIC$ scores where the linear mixed effects model that incorporates patient-specific CD4 variability is used for the longitudinal submodel and Weibull model used for survival submodel are joined by taking different forms of the latent processes $W_{i1}(s)$ and $W_{i2}(t)$. The simple joint models $M_1$ and $M_2$ with no random effects for longitudinal submodel is fitted first, which have a large (poor) total DIC. Next, random intercepts are introduced in the longitudinal submodel. The incorporation of random intercepts in the longitudinal submodel improves $DIC_1$ and also the total $DIC$. Models $M3$ to $M6$ include random intercepts in $W_{i1}(s)$, which results in high improvement in $DIC_1$ for the longitudinal submodel and the total $DIC$ scores. Then, different latent associations through the random intercepts and random variances are introduced. Models $M7$ to $M12$ have both random intercepts and slopes in the longitudinal submodel which results in a substantial decrement in $DIC_1$. But, the incorporation of a frailty term, $U_{3i}$, in $W_{i2}(t)$ increased the value of $DIC_2$ in general as compared
to models which does not include frailty term. Hence, the inclusion of frailty term does not seem to improve the total DIC at all.

Table 3: Results for Joint Model Selection

<table>
<thead>
<tr>
<th>Model</th>
<th>$W_1(t)$</th>
<th>$W_2(t)$</th>
<th>DIC1</th>
<th>DIC2</th>
<th>$d_{Total}$</th>
<th>$pD_{Total}$</th>
<th>DIC_Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1$</td>
<td>0</td>
<td>0</td>
<td>14171.800</td>
<td>3413.350</td>
<td>17364.900</td>
<td>220.263</td>
<td>17583.200</td>
</tr>
<tr>
<td>$M_2$</td>
<td>$t_1$</td>
<td>$t_1$</td>
<td>14172.150</td>
<td>3413.960</td>
<td>17384.310</td>
<td>220.890</td>
<td>17605.200</td>
</tr>
<tr>
<td>$M_3$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12796.200</td>
<td>3413.370</td>
<td>17369.570</td>
<td>408.860</td>
<td>16208.800</td>
</tr>
<tr>
<td>$M_4$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12790.200</td>
<td>3415.330</td>
<td>17364.930</td>
<td>226.141</td>
<td>17582.000</td>
</tr>
<tr>
<td>$M_5$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12789.700</td>
<td>3413.350</td>
<td>17363.050</td>
<td>340.804</td>
<td>16206.000</td>
</tr>
<tr>
<td>$M_6$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12359.300</td>
<td>3413.350</td>
<td>15139.650</td>
<td>633.365</td>
<td>15773.000</td>
</tr>
<tr>
<td>$M_7$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12359.300</td>
<td>3415.290</td>
<td>15141.190</td>
<td>633.503</td>
<td>15774.600</td>
</tr>
<tr>
<td>$M_8$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12359.300</td>
<td>3415.290</td>
<td>15141.190</td>
<td>633.503</td>
<td>15774.600</td>
</tr>
<tr>
<td>$M_9$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12360.900</td>
<td>3417.372</td>
<td>15142.872</td>
<td>635.372</td>
<td>15778.100</td>
</tr>
<tr>
<td>$M_{10}$</td>
<td>$t_1$</td>
<td>$2_t_1$</td>
<td>12360.900</td>
<td>3417.372</td>
<td>15142.872</td>
<td>635.372</td>
<td>15778.100</td>
</tr>
</tbody>
</table>

Generally, Model $M_9$ emerges with the smallest effective number of parameters (less complex or more parsimonious model) among the candidate models. Model $M_7$ has the smallest total DIC (fits the data well) among all other models. Since $W_2(t) = 0$ in model $M_7$, the data set used for this paper does not support the use of joint model to relate a patients survival time to the characteristics driving the patients longitudinal data pattern. This is clinically not reasonable, since high CD4 count represents better health status; patients with CD4 counts that are low or more rapid decline would be expected to have poorer survival. As it is evident from the output of the joint model $M_9$, the use of joint model is apparently not justified for these data, as indicated by the increase in the DIC score and the insignificance of the association parameter $\gamma_2$ with (95% posterior credible interval (-0.174,0.161)) that include zero.

The posterior estimates of the regression coefficients $\beta_1$ and $\beta_2$ with their 95% confidence intervals for final joint model $M_9$ are summarized in Table 4. Here the results in both separate and joint analyses are the same for longitudinal data. In the longitudinal submodel linear and quadratic time, sex and tobacco addiction are statistically significant at 0.05 level of significance (in the Bayesian sense; 95%credible set excludes 0), while knowledge of ART and condom use are statistically significant at this level in the survival submodel. However, the posterior estimates of the association parameter $\gamma_2$ in the joint analysis is insignificant, providing very weak/No evidence of association between the two submodels and indicating that both the initial level and slope of CD4 count is not associated with the hazard of death.
### Table 4: Analysis of Final Joint Model $M_9$

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Posterior Mean</th>
<th>st.dev</th>
<th>MC error</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Longitudinal Submodel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\beta_{11}$)</td>
<td>13.670</td>
<td>0.3136</td>
<td>0.0060</td>
<td>(13.050, 14.280)</td>
</tr>
<tr>
<td>Obstime ($\beta_{12}$)</td>
<td>2.151</td>
<td>0.0836</td>
<td>0.0009</td>
<td>(1.987, 2.314)</td>
</tr>
<tr>
<td>(Obstime)$^2$ ($\beta_{13}$)</td>
<td>-0.118</td>
<td>0.0096</td>
<td>0.0001</td>
<td>(-0.137, -0.099)</td>
</tr>
<tr>
<td>Sex ($\beta_{14}$)</td>
<td>-0.932</td>
<td>0.386</td>
<td>0.0080</td>
<td>(-1.679, -0.167)</td>
</tr>
<tr>
<td>Functional ($\beta_{15}$)</td>
<td>-0.417</td>
<td>0.3423</td>
<td>0.0065</td>
<td>(-1.089, 0.250)</td>
</tr>
<tr>
<td>Alcohol ($\beta_{16}$)</td>
<td>-0.959</td>
<td>0.5559</td>
<td>0.0120</td>
<td>(-2.057, 0.137)</td>
</tr>
<tr>
<td>Tobacco ($\beta_{17}$)</td>
<td>1.148</td>
<td>0.5640</td>
<td>0.0122</td>
<td>(0.038, 2.270)</td>
</tr>
<tr>
<td>OIS ($\beta_{18}$)</td>
<td>0.053</td>
<td>0.0873</td>
<td>0.0018</td>
<td>(-0.120, 0.222)</td>
</tr>
<tr>
<td>tau1</td>
<td>0.106</td>
<td>0.0100</td>
<td>0.0001</td>
<td>(0.087, 0.128)</td>
</tr>
<tr>
<td>tau2</td>
<td>2.398</td>
<td>0.3469</td>
<td>0.0048</td>
<td>(1.802, 3.165)</td>
</tr>
<tr>
<td>$\mu_v$</td>
<td>2.033</td>
<td>0.0530</td>
<td>0.0007</td>
<td>(1.929, 2.136)</td>
</tr>
<tr>
<td>$\tau_v$</td>
<td>2.143</td>
<td>0.3346</td>
<td>0.0065</td>
<td>(1.588, 2.891)</td>
</tr>
<tr>
<td>Survival submodel</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ($\beta_{21}$)</td>
<td>-4.010</td>
<td>0.1456</td>
<td>0.0019</td>
<td>(-4.298, -3.727)</td>
</tr>
<tr>
<td>TB status ($\beta_{22}$)</td>
<td>0.266</td>
<td>0.1398</td>
<td>0.0008</td>
<td>(-0.012, 0.335)</td>
</tr>
<tr>
<td>Knowledge ($\beta_{23}$)</td>
<td>-0.257</td>
<td>0.0717</td>
<td>0.0006</td>
<td>(-0.399, -0.117)</td>
</tr>
<tr>
<td>Condomuse ($\beta_{24}$)</td>
<td>0.347</td>
<td>0.1321</td>
<td>0.0014</td>
<td>(0.089, 0.607)</td>
</tr>
<tr>
<td>OIS ($\beta_{25}$)</td>
<td>-0.008</td>
<td>0.026</td>
<td>0.0002</td>
<td>(-0.060, 0.042)</td>
</tr>
<tr>
<td>Living room ($\beta_{26}$)</td>
<td>-0.016</td>
<td>0.0476</td>
<td>0.0005</td>
<td>(-0.112, 0.075)</td>
</tr>
<tr>
<td>$\gamma_2$ (Ascc. parameter)</td>
<td>-0.006</td>
<td>0.0850</td>
<td>0.0004</td>
<td>(-0.174, 0.161)</td>
</tr>
<tr>
<td>$p$ (shape parameter)</td>
<td>3.995</td>
<td>0.1633</td>
<td>0.0070</td>
<td>(3.692, 4.346)</td>
</tr>
</tbody>
</table>

When evaluating the overall performance of both the separate and joint models in terms of model goodness of fit, the separate model performs better. But, joint model is found to be better in terms of the effective number of parameters. The effective number of parameters of the separate $M_7$ and joint model $M_9$ are 633.365 and 633.078, respectively, while the posterior means of the deviance functions are 15139.60 and 15141.50. As a result, the corresponding $DICs$ for the separate and joint models are 15773 and 15774.600, respectively. Hence, posterior mean of the deviance function of the separate model is smaller, which results in smaller total $DIC$ score, than that of the joint model. Therefore, the separate model fits the data better than joint model $M_9$.

Regarding to the submodels, the $DIC$s of the longitudinal submodel in the separate and joint models are 12359.60 and 12359.30, respectively, which is somewhat larger in the separate model. But the respective $DIC$s of the survival submodel in the separate and joint models are 3413.350 and 3415.290, the survival submodel has smaller $DIC$ in the separate model. In general, the separate model is preferred as it has a smaller total $DIC$ than the joint model. The statistical insignificance of the association parameter $\gamma_1$ is also another evidence that the separate model is better than the joint model.

### Assessing Chain Convergence

In all of the joint models, three parallel MCMC sampling chains, 50,000 iteration each and 25,000 burn-in period, with different starting values are used. One of the initial values is obtained from the separate analysis, the other is by randomly selecting from the corresponding prior distributions and the third one is set to be null for all param-
eters.

Time series plot of the history of iterations of the final joint model and separate model shows a reasonable degree of randomness between iterations and also the overlaps of the three chains indicates that the same solutions are obtained for each initial values. Therefore, the Gibbs sampler has been converged to the target density. Moreover, MC error can be checked. Since the values of MC errors are very low in comparison to its posterior summaries (especially its standard error) then the posterior density has converged to target density.

4 Conclusions

The objective of this study was to investigate the Bayesian joint model of longitudinal CD4 measurements and time-to-death event of HIV/AIDS patients. The patients had been under ART follow-up of at ShaShemene referral hospital. The method includes shared random effects which induce association between the longitudinal data and the survival outcome by incorporating subject specific variances which possesses some attractive features on modeling longitudinal response. Both separate and joint analysis are conducted. The separate analysis is preferred for several reasons. Firstly, it helps to specify the mean response of the model. Secondly, the random effects to be included in the longitudinal model can be easily determined, and thirdly initial values to be provided for the joint models can be obtained.

In the separate analysis of the longitudinal data, the CD4 measurements are checked for normality using normal plot with histogram. The plots indicates that there is a deviation from normality and needs some transformation. After a square root transformation of the CD4 counts, the mean response of the longitudinal square root CD4 counts is determined to be quadratic in time. Then, the data are analyzed using the linear mixed effects model incorporating patient specific square root CD4 variability. The patient specific variability is significant which supports the assumption of heterogeneous variances. In the longitudinal submodel, linear and quadratic time, sex and tobacco addiction are statistically significant at 0.05 level of significance. Out of the covariates included in the survival submodel of the joint model, knowledge of ART and condom use are found to be significantly associated with time to death at 0.05 level of significance. Also, the results show that the patients’ survival in the ART treatment is not significantly associated with patient specific CD4 fluctuations.

For the separate analysis of the survival data, the variables to be included in the survival model are determined using backward LR variable selection method using. Then, of the covariates considered, only five of them: TB status, knowledge of ART, condom use, number of living room, and number of opportunistic infections were extracted to be included. Weibull model expresses both proportional hazards and accelerated failure time models. Out of the covariates included in the survival submodel of the joint model, knowledge of ART and condom use are found to be significantly associated with time to death at 0.05 level of significance.
The Bayesian analysis of joint models with a variety of latent processes are investigated. First, a simple joint model with no random effects in both submodels is fitted and then other 11 models with different random effects and various latent associations of the two submodels are investigated. The results showed that the separate model was found to be statistically significant while joint models were not. The Bayesian joint model is insignificant at 5% significance level for the data considered in this study. It can be concluded that the statistical results obtained from the separate analyses are consistent with those obtained from the joint model. So the joint model is still recommended to apply to such longitudinal and survival data sets.

5 References


