MULTILEVEL JOINT ANALYSIS OF LONGITUDINAL AND BINARY OUTCOMES

by

Seo Yeon Hong

BA in Economics, Ewha Womans University, South Korea, 2000

MA in Economics, Ewha Womans University, South Korea, 2002

MA in Economics, Rutgers, The State University of New Jersey, 2005

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Seo Yeon Hong

It was defended on

November 19, 2012

and approved by

Lisa A. Weissfeld, Ph.D., Professor,

Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Jong-Hyeon Jeong, Ph.D., Associate Professor,

Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Chung-Chou H. Chang, Ph.D., Associate Professor,

Departments of Medicine and Biostatistics, School of Medicine and Graduate School of Public Health, University of Pittsburgh

Matthew Rosengart, M.D., MPH, Associate Professor,

Department of Surgery, University of Pittsburgh

Dissertation Director: **Lisa A. Weissfeld, Ph.D.**, Professor,

Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

MULTILEVEL JOINT ANALYSIS OF LONGITUDINAL AND BINARY OUTCOMES

Seo Yeon Hong, PhD

University of Pittsburgh, 2012

Joint modeling has become a topic of great interest in recent years. The models are simultaneously analyzed using a shared random effect that is common across the two components. While these methods are useful when time-to-event data are available, there are many cases where the outcome of interest is binary and a logistic regression model is used. We propose the use of a joint model with a logistic regression model being used for the binary outcome and a hierarchical mixed effects model being used for the longitudinal outcome. We link the two sub-models using both subject and cluster level random effects and compare it with models using only one level of random effects. We use the Gaussian quadrature technique implemented in the software package aML (Multiprocess Multilevel Modeling software). Simulation studies are presented to illustrate the properties of the proposed model. We also applied our model to the repeated measures of mid-arm muscle circumference (MAMC) and mortality rate for patients within 75 units from 15 centers from a randomized study of hemodialysis (HEMO) and found that the model performs well. We further extend this work by developing methods that can be used to calculate individualized predictions based on our proposed joint model. We use the Bayesian approach to obtain these predictions and implement the method in the software package WinBUGS. The proposed method provides a mechanism for understanding the relationship between a longitudinal measure and a given binary outcome. Thus, it can be used to address several types of public health problems. First, it can be used to understand how changes in a biomarker or other longitudinal measure are related to changes in status of a subject. Second, it can be used to predict the outcome

of a subject based on the trajectory of the longitudinal outcome providing information that can be used in a personalized medicine setting. This allows researchers to identify potentially harmful patterns and intervene at an earlier stage.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to express my thanks and appreciation to all those who helped me in completing my Ph.D. study.

First, I would like to express my appreciation to my dissertation advisor, Dr. Lisa Weissfeld, for her assistance, support, patience and guidance. Also, I want to thank my committee members, Dr. Jong-Hyeon Jeong, Dr. Chung-Chou H. Chang and Dr. Matthew Rosengart for their valuable motivations, suggestions and comments. I also would like to thank to Dr. Douglas White from Department of Critical Care Medicine for the financial support. I would like to thank Dr. Mark Unruh, for access to the HEMO data.

Finally, I would like to thank God for the grace to finish this research and my family; mom, dad, parents-in-law, for their continuous pray and support. Especially, I would like to extend my apprecation to my husband, Hakkyun, and my little daughter, Linsey, for all their support and love.

1.0 LITERATURE REVIEW

1.1 JOINT MODEL

Joint modeling has become a topic of great interest in recent years. The focus has been primarily on models that include a longitudinal and a survival component. In the past, the repeated measures and failure outcomes are treated separately. However, when the two outcomes are related, the separate models may be both inappropriate and uninformative. Joint models have several advantages. They can incorporate (i) time-varying variables, (ii) informative dropout and (iii) censored longitudinal covariates.

The most common approach to the construction of a joint model is the use of shared random effects. The models are simultaneously analyzed using a shared random effect that is common across the two components. The random effects, b_i, which usually follow a normal distribution, are used to link the repeated measures, y_i, and the event time of survival outcome, d_i. The joint likelihood is as follows [1]:

$$\begin{aligned} \mathrm{L} &= \prod_{i=1}^{n} \int f(y_i, d_i|b_i) f(b_i) db_i \\ &= \prod_{i=1}^{n} \int f(y_i|b_i) f(d_i|b_i) f(b_i) db_i. \end{aligned}$$

where $f(b_i)$ has a certain distribution.

We will first review the joint model for a longitudinal outcome and a survival outcome. Then, we will summarize the joint modeling of a longitudinal and a binary outcome. Finally, we will assess the joint model in multilevel data.

1.1.1 Joint Model of Longitudinal and Survival Data

Tsiatis *et al.* [2] used a two-stage approach where the longitudinal process and the Cox proportional hazards model is incorporated. In the first stage, the linear mixed effects model is used and then in the second stage, the empirical Bayes estimates replaced the covariates of the Cox model. They also applied the method to survival and CD4 counts in patients with AIDS.

Schluchter [3] discussed several methods for modeling informatively censored longitudinal data. When informative censoring occurs, it causes problems such as bias in the standard likelihood-based analyses. He pointed out that the Expectation-Maximization (EM) algorithm has several advantages when applied to this problem, namely that it allows for unbalanced data and unequally spaced time intervals. The EM algorithm (Dempster, Laird and Rubin [4]) has been a popular method in the joint random effects model framework. The EM algorithm is an iterative optimization method used to estimate some unknown parameters. More generally, the EM algorithm can be described as follows for a general likelihood function;

$$\begin{aligned} l(\theta) = logp(x|\theta) &= log \sum_z p(x,z|\theta) \\ &= log \sum_z q(z|x,\theta) \frac{p(x,z|\theta)}{q(z|x,\theta)} \\ &\geq \sum_x q(z|x,\theta) log \frac{p(x,z|\theta)}{q(z|x,\theta)} \equiv F(q,\theta), \end{aligned}$$

where X denotes the observed variables, Z denotes the unobserved latent variables and θ is an unknown parameter vector. Note that $logp(x,z|\theta)$ is a complete log-likelihood and $q(z|x,\theta)$ is an arbitrary density over Z. In the E-step, the random coefficients are estimated using the conditional expectation of the log-likelihood given the observed data, that is,

$$\textbf{E-step}: \; q^{(t+1)} = \underset{q}{\operatorname{argmax}} \, F(q, \theta^{(t)}).$$

Then, in the M-step, the expected log-likelihood is maximized using the estimated random effects yielding

$$\textbf{M-step}: \ \theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}} \, F(q^{(t+1)}, \theta).$$

DeGruttola and Tu [5] applied the EM algorithm to a joint model of the progression of CD4-lymphocyte count and survival time using random effects. They assumed that the joint distribution of the covariate process and survival times were multivariate normal. Wulfshon and Tsiatis [6] relaxed some of the assumptions placed on prior methods. They assumed that the individual's random effects need not be normally distributed at every time point since the risk set changes at each failure time. Henderson *et al.* [7] proposed that the two models are linked by a common latent stochastic process. The researchers used a latent bivariate Gaussian process $W(t) = \{W_1(t), W_2(t)\}$ and assumed that the repeated measurement and the failure time process were conditionally independent given $W(t)$ and the covariates. The association of the two outcomes was created through the cross-correlation between $W_1(t)$ and $W_2(t)$. Therefore, the joint distribution of both outcomes for the ith subject is assumed to be an unobserved or latent zero-mean bivariate Gaussian process, $W_i(t) = \{W_{1i}(t), W_{2i}(t)\}$. The sub-model for the repeated measurements part is as follows;

$$\mathrm{Y}_{ij} = X_{1i}(t)'\beta_1 + W_{1i}(t_{ij}) + \epsilon_{ij}, \tag{1.1}$$

where $X_{1i}(t)$ are possiblely time-varying covariates and β_1 are the corresponding coefficients associated with the repeated measures. The sequence of mutually independent measurements errors, ϵ_{ij} is assumed to be $\mathrm{N}(0, \sigma_\epsilon^2)$. The other sub-model is the semi-parametric model for the event intensity process at time t, and is given by

$$\lambda_i(t) = H_i(t)\lambda_0(t)exp\{X_{2i}(t)'\beta_2 + W_{2i}(t)\}, \tag{1.2}$$

with $H_i(t)$ being the zero-one process and $\lambda_0(t)$ being the unspecified form of the baseline hazard. Estimation for the models described in (1.1) and (1.2) is based on an extension of the EM algorithm. Lin *et al.* [8] introduced a joint model comprised of multiple longitudinal outcome variables and a survival outcome. The model incorporated the correlations among longitudinal co-variables and used a one-step-late EM algorithm.

Song *et al.* [9] also tested the sensitivity of violation of the normality assumption in a joint model with a longitudinal and a survival outcome based on the EM algorithm. They relaxed the assumption of a smooth density and found that the results were remarkably robust and consistent to the assumption of normality. For the same joint model, Hsieh *et al.* [10] confirmed that the joint likelihood with the normality assumption of random effects is robust and efficient as long as the longitudinal outcome does not carry large measurement errors.

The Markov chain Monte Carlo (MCMC) methods are another widely used algorithm in the joint model setting. In Bayesian inference, the unknown random effects are estimated based on the posterior distribution. The previous sample values are used to randomly generate the next set of sample values until the Markov chain reaches a stationary distribution. Faucett and Thomas [11] simultaneously analyzed repeated measurements as a prediction of disease risk. They used the Markov chain Monte Carlo technique of Gibbs sampling to estimate the unknown parameters. Gibbs sampling is a special case of the Metropolis-Hastings algorithm. Suppose that $x = (x_1, x_2, \ldots, x_k)$ is from a joint distribution $p(x_1, x_2, \ldots, x_k)$. The samples are updated as follows:

$$
\begin{aligned}
X_1^{(t+1)} &\sim p(x_1 | x_2^{(t)}, x_3^{(t)}, \ldots, x_k^{(t)}) \\
X_2^{(t+1)} &\sim p(x_2 | x_1^{(t+1)}, x_3^{(t)}, \ldots, x_k^{(t)}) \\
X_3^{(t+1)} &\sim p(x_3 | x_1^{(t+1)}, x_2^{(t+1)}, x_4^{(t)} \ldots) \\
&\vdots \quad \vdots \\
X_k^{(t+1)} &\sim p(x_k | x_1^{(t+1)}, x_2^{(t+1)}, \ldots, x_{k-1}^{(t+1)}).
\end{aligned}
$$

From the updated samples, the most recent values are used. Their method reduced bias in the parameter estimates due to covariate measurement error and informative censoring. Later, Xu and Zeger [12] generalized the Markov chain Monte Carlo technique using a latent variable. Guo and Carlin [13] developed the method of Henderson *et al.* into a fully Bayesian

version using the Markov chain Monte Carlo technique. The researchers applied the method to the CD4 counts and time to death simultaneously via a latent bivariate Gaussian process.

Hogan and Laird [14] focused on incomplete data in the longitudinal measurements in the joint modeling setting. They used the missing mechanisms described in Little and Rubin [15], and discussed below. If missingness does not depend on the data Y, then it is defined as missing completely at random (MCAR). In this case

$$f(M|Y,\phi) = f(M|\phi) \text{ for all } Y,\phi,$$

where Y is the complete data, M is missing-data indicator matrix and ϕ denotes the unknown parameters. If missingness depends only on observed components, Y_{obs}, then the missing-data mechanism is called missing at random (MAR) and

$$f(M|Y,\phi) = f(M|Y_{obs},\phi) \text{ for all } Y_{mis},\phi,$$

where Y_{obs} denotes the observed components and Y_{mis} denotes the missing components. If missingness depends on the missing values in the data Y, then it is called not missing at random (NMAR). They also discussed selection and mixture models for the evaluation of missing values in both longitudinal and survival outcomes.

Tsiatis and Davidian([16], [17]) generalized the joint model by placing no assumptions on the distribution of the random effects. Prior to this work, most methods assumed that the random effects followed a normal distribution. They relaxed the assumptions of normality of the random effects and their semiparametric approach may cause a loss of efficiency relative to the models where the parametric specification for the random effects is made. They used the conditional score (CS) approach proposed by Stefanski and Carroll [18]. The conditional score is obtained by conditioning on certain sufficient statistics when the explanatory variables are fixed constants. Later, Song *et al.* [19] extended this approach to include multiple, possibly correlated, time-dependent covariates using the conditional score method.

Proc NLMIXED [20] is an another convenient tool for estimation in the joint modeling setting. Vonesh *et al.* [21] proposed a joint model comprised of parametric or semiparametric survival models for the survival component and the generalized linear or non-linear mixed-effects models for the longitudinal component. This allows flexibility in the specification of

the distribution of the event time and the association of the longitudinal outcome. They used the Laplace approximation implemented in Proc NLMIXED in SAS to obtain estimates from the model. Liu [22] considered a joint model when the values of the repeated measures contain a large number of zero values and right-skewed positive values. They applied the method to the longitudinal monthly medical costs of chronic heart-failure patients using the Gaussian quadrature techniques implemented in SAS Proc NLMIXED [23].

We have discussed several estimation methods for the joint model of longitudinal and survival outcomes. Due to the unknown random effects, the Expectation-Maximization (EM) and Markov chain Monte Carlo (MCMC) algorithms have been the popular methods and extended in many ways. Also, some papers focused on incomplete data in longitudinal measurements. The conditional score (CS) approach was another method used to relax the normality assumption of random effects. Recently, Proc NLMIXED in SAS has been conveniently utilized, since it can easily accommodate a user-specified likelihood. While these methods are useful when time-to-event data are available, there are many cases where the outcome of interest is binary and a logistic regression model is used. We will discuss some papers where a binary outcome was used in place of a survival outcome.

1.1.2 Joint Model of Longitudinal and Binary Data

Wang *et al.* [24] considered a joint model using a naive estimator. They used a regression calibration (RC) estimator, where repeated measurements were assumed to be linear. However, when the measurements were nonlinear, the RC model produced a biased estimator. They used the refined RC (RR) estimator in logistic regression to generate an unbiased estimate. Li *et al.* [25] introduced the sufficiency score (SS) and the conditional score (CS), which generalized the linear model with no distributional assumptions on the random effects and consistent inference even when the distribution was misspecified. Later, Li *et al.* [26] developed the generalized sufficiency score (GSS) and the generalized conditional score (GCS), which have no distributional nor covariance structural assumptions placed on the covariate random effects. These approaches are more flexible and are applicable to the multivariate longitudinal covariate processes.

Horrocks *et al.* [27] considered the prediction of pregnancy based on longitudinal measurements of adhesiveness of certain blood lymphocytes. They used the joint model developed by Wang *et al.* [24] and used a Bayesian approach to estimate the parameters as implemented in WinBUGS. Wannemuehler *et al.* [28] explored the joint model of longitudinal and logistic health outcomes in left- and interval-censored data. They compared the joint model with a two-stage approach, and the joint modeling approach implemented in SAS Proc NLMIXED performed better with little bias and near-nominal confidence interval coverage.

In summary, the sufficiency score (SS) and the conditional score (CS) methods were used to relax the distributional assumptions on the random effects. Also, Bayesian methods and the likelihood-based approach were discussed in the joint modeling of a longitudinal and a binary outcome.

1.1.3 Joint Model with Multilevel Data

Hierarchical or clustered structures are presnet in many biomedical studies. A common example is in longitudinal studies where an individual's repeated measures are correlated with each other. Individuals may be further nested within geographical areas or institutions such as schools or hospitals. In this situation, there are clustering effects due to the correlation between the same levels nested within the higher levels. As a result of this clustering, there are different random effects between levels, so we can use mixed model to account for these random effects.

Ratcliffe *et al.* [29] developed a joint model for the longitudinal and survival outcome in multilevel data. They linked the two submodels with cluster-level random effects and applied the EM algorithm to estimate the parameters. In the presence of clustered data, the cluster-level linkage performed better than a subject-level link.

Liu *et al.* [30] linked the joint model of a longitudinal and a survival outcome using both cluster level and subject level random effects. They used the Gaussian quadrature technique implemented in aML (Multiprocess Multilevel Modeling software [31]). The researchers showed that failure to satisfy the assumptions placed on the dependence structure between the outcomes can cause serious biases.

The joint model of longitudinal and binary outcomes in multilevel data is still under development. The focus of this paper is on the setting where we only have information about the events (dead or alive), but not the failure time in the presence of the clustering effect. This model is applicable in a setting where the data are hierarchically structured, such as, hospitals at level 1, patients at level 2, and longitudinal measurements at level 3. In other words, when the subjects are clustered within higher levels, the random effects are different between their levels. This work focuses on the setting where the longitudinal outcome is subject to multiple hierarchical levels and the second outcome is binary. We propose the use of a joint model with a logistic regression model being used for the binary outcome and a hierarchical mixed effects model being used for the longitudinal outcome. Here, we link the two sub-models using both subject and cluster level random effects and compare it with models using only one level of random effects. We present simulation results that compare the results of a single level random effect to the multi-level random effects. Then we apply this model to the motivating example of the HEMO data [32].

Another major issue in joint modeling is the estimation of predicted values for each of the observations in the data set. While there has been some research related to prediction for the joint model of longitudinal and survival outcomes, there has been little work for joint model of longitudinal and binary outcomes. To address this issue we obtain individual predictions for our proposed joint model. Through fitting our model in WinBUGS, which is based on the Bayesian method using a Markov chain Monte Carlo (MCMC), we will obtain estimates of the individual predicted probability. Then, we can assess the performance of the longitudinal measures in the prediction of mortality.

2.0 MULTILEVEL JOINT ANALYSIS OF LONGITUDINAL AND BINARY OUTCOME

2.1 INTRODUCTION

Joint modeling has become a topic of great interest in recent years, seeing extensions to methodology for jointly modeling two different outcomes, extensions to accommodate informative censoring and methods for incorporating censored longitudinal covariates into a survival model. These methods address the need to link longitudinal information to other types of models and provide for improved estimates. This is in contrast to traditional methods where each component is modeled separately or where the longitudinal component is included as a time-varying covariate in a survival model, providing potentially biased results. While joint models offer the advantage of allowing for the joint assessment of two outcomes and the reduction in bias in estimates, the models can be difficult to fit and need extra care in interpretation.

The focus of work in the area of joint modeling has primarily been on the development of models that include a longitudinal and a survival component in a single model. Research in this area has centered on several different approaches; the use of the expectation-maximization (EM) and Markov chain Monte Carlo (MCMC) algorithms to obtain estimates for the joint model and the use of the full likelihood. The EM algorithm has been used by many researchers. DeGruttola and Tu [5] assumed that the joint distribution of the covariate process and survival times were multivariate normal. Wulfshon and Tsiatis [6] assumed that the individual's random effects need not be normally distributed all the time when the individuals are at risk at each event time and thus removed from observation. Henderson *et al.* [7] proposed that the two models are linked by a common latent stochastic process. Lin

et al. [8] introduced multiple longitudinal variables and the survival outcome. The model incorporated the correlations among longitudinal co-variables. Song *et al.* [9] relaxed the assumption with a smooth density and found out the results were remarkably robust and consistent to the assumption of normality. Hsieh *et al.* [10] confirmed that the joint likelihood with normality assumption of random effects is robust and efficient as long as the longitudinal outcome does not carry large measurement errors.

Several researchers applied the MCMC approach to joint modeling as well. Faucett and Thomas [11] used the Markov chain Monte Carlo technique of Gibbs sampling to estimate unknown parameters. Xu and Zeger [12] generalized the Markov chain Monte Carlo technique using a latent variable. Guo and Carlin [13] developed the method of Henderson into a fully Bayesian version using the Markov chain Monte Carlo technique.

PROC NLMIXED in SAS [20] has been another widely used piece of statistical software for joint modeling using the likelihood approach. Vonesh *et al.* [21] proposed a joint model including parametric and semiparametric survival models and generalized linear or nonlinear mixed-effects models which allow for flexibility in the specification of the distribution of the event time for the survival component of the joint model. They used the Laplace approximation implemented in PROC NLMIXED in SAS for fitting this model. Liu [22] also used PROC NLMIXED for the case where the repeated measures outcome contains a large number of zero values and right-skewed positive values. Tsiatis and Davidian ([16], [17]) generalized the joint model by developing a model where there are no assumptions placed on the distribution of the random effects. They used the conditional score (CS) approach proposed by Stefanski and Carroll [18]. This was later extended to the setting where there are multiple, possibly correlated, time-dependent covariates (Song *et al.* [19]).

The motivation for this work arose from several practical applications where one of the outcomes of interest was a binary outcome, rather than a survival outcome. One example arose out of a study of intensive care units looking at the relationship between in-hospital mortality and daily sunlight exposure while in the intensive care unit. The second example came out of a clinical trial of dialysis approaches and mortality. In these setting, there was interest in examining the relationship between longitudinal measures of mid-arm muscle circumference (MAMC) and mortality. While there has been some development of joint models

for binary and longitudinal outcomes, these modeling techniques are not as well developed as those that are available for a joint model including survival and a longitudinal outcome and there have been no models developed for hierarchically structured data. Horrocks [27] used a Bayesian approach for the prediction of pregnancy and other authors have used the sufficiency score (SS) and the conditional score (CS) to develop joint models with a binary and longitudinal outcome with no assumptions placed on the random effects ([25], [26]). The incorporation of multilevel data into the joint model has not been as widely studied. Ratcliffe *et al.* [29] developed a joint model for a longitudinal and survival outcome in multilevel data linking the two submodels with cluster-level random effects. The EM algorithm was then used to obtain parameter estimates from this joint model. They then showed that cluster-level linkage performed better than a subject-level linkage in the presence of clustered data. Liu *et al.* [30] linked the joint model composed of a longitudinal and survival outcome using both cluster level and subject level random effects. To implement this method, they used the Gaussian quadrature technique implemented in aML (Multiprocess Multilevel Modeling software [31]). They also showed that the model is sensitive to the assumptions on the dependence structure between the outcomes, resulting in serious bias when these assumptions are violated.

The focus of this work is on the development of a joint model in the multilevel data setting where the outcomes include a binary and a longitudinal outcome. This model is applicable in setting where data are hierarchically structured, such as, hospitals at level 1, patients at level 2, and longitudinal measurements at level 3. In other words, when the subjects are clustered within higher levels, the random effects are different between their levels. This work focuses on the setting where the longitudinal outcome is subject to multiple hierarchical levels and the second outcome is binary. We propose the use of a joint model with a logistic regression model being used for the binary outcome and a hierarchical mixed effects model being used for the longitudinal outcome. Here, we link the two sub-models using both subject and cluster level random effects and compare it with models using only one level random effects. We showed the simulation results which were compared with the results of one level random effects. Then we applied this model to the motivating example of the HEMO data [32].

2.2 METHODS

Let Y_{ijk} denote the k-th repeated measure for the j-th subject within i-th cluster ($i = 1, 2, ..., n, j = 1, 2, ..., n_i,$ and $k = 1, 2, ..., m_{ij}$). Let p_{ij} denote the probability of response for the j-th subject of a binary zero-one outcome variable, X_{ij}. Let Z_{ijk} and W_{ij} denote the covariate vectors of fixed effects for the longitudinal outcome model and the logistic regression model, respectively. We assume that the missing mechanism is missing at random (MAR). Under this assumption, we keep all of the observations in the sample whether the observations are missing or not. The missing data simply do not contribute to the estimation. As shown in (2.2) below, the repeated measures k are indexed from 1 to m_{ij}. In this model, a_i and b_{ij} denote the random effects at the cluster and subject levels, respectively. They are assumed to be independent and identically distributed according to a normal distribution with mean 0 and corresponding variances, e.g. σ_a^2 and σ_b^2.

We define the joint model as

$$\begin{aligned} y_{ijk} &= Z_{ijk}^T\beta + a_i + b_{ij} + e_{ijk} \\ logit(p_{ij}) &= W_{ij}\alpha + \gamma_1 a_i + \gamma_2 b_{ij}, \end{aligned} \tag{2.1}$$

where β and α are unknown vectors of parameters. In this model γ_1 and γ_2 represent the association between the two models at each cluster and subject level. The error term, e_{ijk} is assumed to be N(0, σ_e^2) and independent of (a_i, b_{ij}, X_{ij}). Let O_{ij} denote the observed data for the i-th subject within the j-th cluster, then the likelihood is as follows:

$$L = \prod_{i=1}^{n} L_i = \prod_{i=1}^{n} \int \prod_{j=1}^{n_i} L_{ij}\phi(a_i)da_i \tag{2.2}$$

where

$$L_{ij} = \int L(O_{ij}|a_i, b_{ij})db_{ij}$$

$$= \int \prod_{k=1}^{m_{ij}} \phi(e_{ijk}) \left[\prod_{j=1}^{n_i} p_{ij}^{x_{ij}} (1-p_{ij})^{(1-x_{ij})} \right] \phi(b_{ij}) db_{ij}.$$

Note that

$$\mathrm{e}_{ijk} = y_{ijk} - Z_{ijk}^T \beta - a_i - b_{ij},$$

and that $\phi(a_i)$ represents a normal density function with mean 0 and variance σ_a^2. Likewise, $\phi(b_{ij})$ is the normal density function with mean 0 and variance σ_b^2. Here, x_{ij} denotes the realization of X_{ij}.

We also considered the two reduced models discussed by Liu *et al.* [30]. The first model we consider, 'reduced model A', restricts the number of random effects related to the binary outcome by including only the cluster-level random effect, a_i, in the model. This results in a model where γ_2 in equation (2.1) is set equal to zero and is written as

$$\begin{aligned} y_{ijk} &= Z_{ijk}^T \beta + a_i + b_{ij} + e_{ijk} \\ logit(p_{ij}) &= W_{ij}\alpha + \gamma_1 a_i. \end{aligned} \tag{2.3}$$

The other model we consider, 'reduced model B', restricts the number of random effects related to the binary outcome by including only the subject-level random effect, b_{ij}, in the model. This results in a model where γ_1 in equation (2.1) is set equal to zero and is written as

$$\begin{aligned} y_{ijk} &= Z_{ijk}^T \beta + a_i + b_{ij} + e_{ijk} \\ logit(p_{ij}) &= W_{ij}\alpha + \gamma_2 b_{ij}. \end{aligned} \tag{2.4}$$

The likelihoods of the reduced models are different in terms of the p_{ij} when compared to our model, which includes both subject and cluster level random effects. Note that the likelihood equations presented in (2.2) will include a single integration.

The integration of the random effects in (2.2) requires the use of special software. Proc NLMIXED in SAS is a convenient tool for the fitting of joint models. However, our data is multilevel data and Proc NLMIXED in SAS can not handle clustered level random effects.

We used the aML software to implement the joint model using Gaussian quadrature techniques. This software is specific for the estimation of multilevel and multiprocess models. It uses a numerical integration algorithm based on Gaussian quadrature techniques. We used five quadrature points for the simulation study and it was enough for this purpose. Also, we used 50 quadrature points for our application data to get more accurate results. Huber-corrected standard error estimates were used for the covariance matrix of the parameter estimates which is robust to heteroscedasticity.

2.3 SIMULATION STUDY

We conducted simulation studies to identify the sensitivity of the estimates when the assumptions of the dependence structure were violated. We compared our full model with two reduced models which included only one of the random effect terms. The 'reduced model A' was only linked via the cluster-level random effect and 'reduced model B' was only linked via the subject-level random effect. We generated data for the simulation study following that outlined in Liu *et al.* [30]. For all simulations we simulated data on 500 subjects with a cluster size of 50, so that there were 10 subjects within each of the clusters. The results presented are based on 600 simulated samples for each scenario. The model below was used for generating data:

$$y_{ijk} = \beta_0 + \beta_1 Z_{ij} + \beta_2 time + a_i + b_{ij} + e_{ijk}$$
$$logit(p_{ij}) = \alpha_0 + \alpha_1 Z_{ij} + \gamma_1 a_i + \gamma_2 b_{ij}.$$

Let y_{ijk} represent the repeated measurement at the integer time k, with k ranging from 1 to 5. We generated the subject-level covariate, Z_{ij}, from a binary distribution with probability 0.5. Coefficent parameters were $\beta = (\beta_0, \beta_1, \beta_2)^T = (-1, -.5, -.2)^T$. The random intercepts a_i and b_{ij} were also included in the model. We assumed $a_i \overset{\text{iid}}{\sim} N(0, \sigma_a^2)$ with $\sigma_a^2 = 1$ and $b_{ij} \overset{\text{iid}}{\sim} N(0, \sigma_b^2)$ with $\sigma_b^2 = 1$. Since the repeated measures from a subject j share the common random effects b_{ij} and the subjects from a cluster share the common random effects a_i, the correlation is induced from the random effects. Here, the covariance structure is

compound symmetry where the variance at all time points is the same and the correlation between any two distinct measurements is the same. The covariance structure is as follows:

$$COV(y_{ijk}) =$$

$$\sigma_{a^2} + \sigma_{b^2} + \sigma_{e^2} \begin{pmatrix} 1 & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} \\ \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & 1 & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} \\ \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & 1 & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} \\ \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & 1 & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} \\ \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & \frac{\sigma_{a^2}+\sigma_{b^2}}{\sigma_{a^2}+\sigma_{b^2}+\sigma_{e^2}} & 1 \end{pmatrix}$$

The error term was $e_{ijk} \overset{\text{iid}}{\sim} N(0, \sigma_e^2)$ with $\sigma_e^2 = 1$. The logistic regression included an intercept α_0 and the subject-level covariate Z_{ij} with coeficient $\alpha = (\alpha_0, \alpha_1)^T = (1.5, -1)^T$. Also, the random effects a_i, at the cluster level, and b_{ij}, at the subject level, were included in the model. In case I, we assumed that the coefficient parameters $\gamma = (\gamma_1, \gamma_2)^T = (1, 1)^T$. Non-adaptive Gaussian quadrature with 5 quadrature points was implemented in aML.

Table 1 presents the simulation results. As expected, the estimates of the proposed model had very small biases and the coverage probability showed that the model performed much better than the other two reduced models. The reduced model A was where the model was linked only via the cluster-level random effect. Likewise, the reduced model B was set to be linked only via the subject-level random effects. The reduced model A showed a lower coverage probability percent in the parameters α_0, α_1 and γ_1 than our model. Also, the estimates were biased for the parameters α_0, α_1 and σ_a in the reduced model B.

In case II, we assumed that the coefficient parameters $\gamma = (\gamma_1, \gamma_2)^T = (1, 0)^T$ and other parameters remained the same in case I. In this case, the reduced model A was the true model. The results can be seen in Table 2. The estimates of the proposed model are still unbiased and the coverage probability percent is reasonable. However, reduced model B showed lower coverage probability for α_0, α_1 and γ_2.

In case III, we also generated the data with $\gamma = (\gamma_1, \gamma_2)^T = (0, 1)^T$, where the reduced model B was the true model. The proposed model still showed unbiased estimates and

reasonable coverage probability. In the reduced model A, α_0, α_1 and γ_1 were poorly estimated with low coverage probability. The results are presented in Table 3.

Table 1: Simulation Results for case I : $\gamma = (\gamma_1, \gamma_2)^T = (1, 1)^T$

	Our model			Reduced model A			Reduced model B		
Parameter	Est	SE	CP Percent	Est	SE	CP Percent	Est	SE	CP Percent
$\beta_0 = -1$	-0.997	0.161	94.33	-0.997	0.161	94.34	-0.998	0.161	94.33
$\beta_1 = -0.5$	-0.503	0.101	94.83	-0.503	0.102	95.01	-0.503	0.103	94.67
$\beta_2 = 0.2$	0.200	0.014	94.50	0.200	0.014	94.51	0.201	0.014	94.50
$\alpha_0 = 1.5$	1.563	0.252	93.50	1.303	0.206	80.17	1.288	0.209	78.33
$\alpha_1 = -1$	-1.040	0.261	94.17	-0.869	0.216	90.83	-0.858	0.229	89.33
$\gamma_1 = 1$	1.022	0.147	93.17	0.934	0.119	85.81			
$\gamma_2 = 1$	1.051	0.168	93.67				0.995	0.135	94.17
$\sigma_a = 1$	0.982	0.108	93.50	0.986	0.107	93.99	0.815	0.101	53.33
$\sigma_b = 1$	0.997	0.039	94.83	0.995	0.039	94.99	1.024	0.043	91.83
$\sigma_e = 1$	1.000	0.016	92.50	1.000	0.016	92.50	1.000	0.016	92.50

CP is the coverage probability of the 95 percent confidence interval.

Table 2: Simulation Results for case II : $\gamma = (\gamma_1, \gamma_2)^T = (1, 0)^T$

	Our model			Reduced model A			Reduced model B		
Parameter	Est	SE	CP Percent	Est	SE	CP Percent	Est	SE	CP Percent
$\beta_0 = -1$	-0.998	0.161	94.17	-0.998	0.161	94.17	-0.997	0.161	94.33
$\beta_1 = -0.5$	-0.502	0.101	95.17	-0.502	0.101	95.00	-0.503	0.101	94.67
$\beta_2 = 0.2$	0.201	0.014	94.50	0.201	0.014	94.50	0.201	0.014	94.50
$\alpha_0 = 1.5$	1.512	0.225	93.83	1.506	0.224	93.67	1.268	0.193	73.83
$\alpha_1 = -1$	-1.013	0.222	95.67	-1.010	0.221	96.17	-0.855	0.201	88.67
$\gamma_1 = 1$	1.014	0.147	93.00	1.010	0.145	93.00			
$\gamma_2 = 0$	-0.009	0.124	95.00				0.096	0.108	85.50
$\sigma_a = 1$	0.983	0.108	93.33	0.983	0.108	93.33	0.966	0.108	91.67
$\sigma_b = 1$	0.998	0.039	95.00	0.998	0.039	95.00	1.000	0.040	95.17
$\sigma_e = 1$	1.000	0.016	92.50	1.000	0.016	92.50	1.000	0.016	92.50

Table 3: Simulation Results for case III : $\gamma = (\gamma_1, \gamma_2)^T = (0, 1)^T$

	Our model			Reduced model A			Reduced model B		
Parameter	Est	SE	CP Percent	Est	SE	CP Percent	Est	SE	CP Percent
$\beta_0 = -1$	-1.002	0.170	94.13	-0.997	0.161	94.17	-0.998	0.161	94.17
$\beta_1 = -0.5$	-0.502	0.111	93.96	-0.503	0.101	94.67	-0.502	0.101	93.83
$\beta_2 = 0.2$	0.201	0.014	94.46	0.201	0.014	94.50	0.201	0.014	94.50
$\alpha_0 = 1.5$	1.523	0.306	93.79	1.262	0.152	64.00	1.440	0.174	92.50
$\alpha_1 = -1$	-1.008	0.319	94.63	-0.842	0.199	85.33	-0.955	0.226	93.00
$\gamma_1 = 0$	0.026	0.617	96.31	0.082	0.103	88.33			
$\gamma_2 = 1$	1.030	0.208	93.29				0.976	0.135	91.83
$\sigma_a = 1$	0.986	0.178	93.46	0.983	0.108	93.33	0.984	0.106	93.17
$\sigma_b = 1$	0.998	0.073	95.13	0.998	0.039	95.00	0.998	0.039	95.33
$\sigma_e = 1$	1.000	0.016	93.29	1.000	0.016	92.50	1.000	0.016	92.50

We also simulated data for both random intercept and random slope effects. The model was as follows:

$$y_{ijk} = \beta_0 + \beta_1 Z_{ij} + \beta_2 time + a_i + b_{1ij} + b_{2ij} time + e_{ijk}$$
$$logit(p_{ij}) = \alpha_0 + \alpha_1 Z_{ij} + \gamma_1 a_i + \gamma_2 b_{1ij} + \gamma_3 b_{2ij}$$

The model was linked with the three random effects, which were the cluster-level random intercept a_i, the subject-level random intercept b_{1ij} and the subject-level random slope b_{2ij}. The b_{2ij} shows the within subject variation over time. The coefficient parameter γ_3 and the variance component σ_{b2}^2 were set to be .5. The results for the proposed model showed reasonable estimates. The results are shown in Table 4.

We also conducted several different simulation studies by changing the sample size and the cluster size. As the sample size gets smaller, we found that the coverage probabilities of the cluster level association and the cluster level variance component were lower. We

also changed the values of the association of the joint model, setting $\gamma_1 = \gamma_2 = 1.5$. These results had lower lower coverage probabilities for the cluster level variance component than the cases where $\gamma_1 = \gamma_2 = 1$.

Table 4: Simulation Results for case IV : $\gamma = (\gamma_1, \gamma_2, \gamma_3)^T = (1, 1, 0.5)^T$

	Our model			Reduced model A			Reduced model B		
Parameter	Est	SE	CP Percent	Est	SE	CP Percent	Est	SE	CP Percent
$\beta_0 = -1$	-0.997	0.166	94.00	-0.997	0.166	94.00	-0.996	0.169	94.89
$\beta_1 = -0.5$	-0.493	0.128	94.33	-0.493	0.128	94.50	-0.495	0.131	94.53
$\beta_2 = 0.2$	0.199	0.034	94.50	0.199	0.034	94.50	0.199	0.034	94.89
$\alpha_0 = 1.5$	1.631	0.278	93.50	1.289	0.205	78.83	1.267	0.204	75.84
$\alpha_1 = -1$	-1.085	0.276	92.33	-0.859	0.213	86.83	-0.848	0.222	85.36
$\gamma_1 = 1$	1.060	0.168	93.67	0.922	0.124	83.83			
$\gamma_2 = 1$	1.119	0.272	97.00				1.026	0.174	94.18
$\gamma_3 = 0.5$	0.532	0.198	92.83				0.415	0.165	89.42
$\sigma_a = 1$	0.984	0.114	92.00	0.991	0.113	92.33	0.803	0.109	52.38
$\sigma_{b_1} = 1$	0.989	0.066	93.83	0.986	0.066	93.83	1.037	0.070	92.42
$\sigma_{b_2} = 0.7$	0.705	0.025	92.50	0.704	0.024	92.83	0.705	0.025	92.42
$\sigma_e = 1$	0.999	0.018	92.17	1.000	0.018	92.00	1.000	0.018	92.42

From the simulation study we presented, the estimates are sensitive to the violation of the assumptions of the dependence structure between the longitudinal outcome and the binary outcome. The results show that the binary part and the variance component have poor coverage probability and biased estimates if we ignore the correct assumptions.

2.4 APPLICATION

The HEMO study was a randomized controlled trial designed to identify the effects of dialysis dose and membrane flux on morbidity and mortality for patients undergoing chronic hemodialysis. The study enrolled 1846 patients nested within 75 units within 15 centers, randomized by dose (standard or high) and by membrane type (high or low). One question of secondary interest is the relationship between mid-arm muscle circumference and overall health. To address this we fit a joint model with the logistic outcome being mortality and the longitudinal component being mid-arm muscle circumference. A mixed model was fit to the data with the longitudinal measure of mid-arm muscle circumference (MAMC) as the outcome [33]. This measure was calculated using the following equation:

$$\text{MAMC (cm)} = \text{mid-arm circumference (cm)}$$
$$- 3.142 \times \text{triceps skinfold (TSF) (cm)}.$$

Our analysis included 1799 patients who had at least one measure of mid-arm muscle circumference (MAMC). Time is calculated from randomization date to the visit date for evaluation in years. The average number of follow-up visits is 3 (range 1-8) and the average follow-up time is 2.09 years (range: 0 - 6.36). Out of 1799 patients, 840 (46.7%) died during the study. The mean MAMC value is 24.6 (range 8.5 - 51.2).

The following baseline variables were included in both models: age at the first visit (mean 58 years), gender (44 % male), Index of Coexisting Disease (ICED) severity score which was calculated with diabetes excluded (36 % with a score of 1, 31 % with a score of 2 and 33 % with a score of 3), diabetic status (44 % diabetes) and race (63 % black). Also, the variables indicating treatment assignment, dose (standard vs. high dose) and membrane type (high or low flux membranes), were considered. Time in years since randomization date was also included in the longitudinal model.

Since patients receive similar treatment if they are in the same unit, there will be a within-cluster correlation for patients who are treated in the same unit (level 1). Also, the repeated measurements of mid-arm muscle circumference (MAMC) (level 3) are nested within a patient (level 2). We assumed that mortality might depend on the repeated measures

of MAMC at both the unit level and the subject level. First, we built the longitudinal submodel with the outcome of MAMC and then we built the logistic regression submodel with the outcome of all-cause mortality. We linked the two sub-models using both subject and cluster level random effects. The results, shown in Table 5, are based on 50 quadrature points.

There was a significant linear decreasing pattern in MAMC over time. It decreased by 0.122 per year with a p-value of <0.001. Patients who had a higher index of coexistent diseases (ICED) score had significantly higher mortality. Also, patients who had diabetes had higher mortality when compared to the patients with no diabetes by 0.317 and a p-value of 0.012. We observed that patients with an older age had a higher mortality by 0.052 with a p-value of <0.001.

We found that the random effects at both the cluster and the subject levels were significant for MAMC. The estimate of σ_a (cluster level) was 0.420 with a p-value of 0.05 and σ_b (subject level) was 3.179 with a p-value <0.001. The longitudinal MAMC was negatively correlated with mortality rate at the subject level (γ_2), but not at the unit level (γ_1).

We also wanted to compare our proposed model with four reduced models. Two of these reduced models include both the cluster level and the subject level of random effects in the mixed model. Reduced model A was only linked via the cluster level. For this model, we found that γ_1 was slightly smaller in magnitude when compared to the full model but it was still not significant. Reduced model B was only linked via the subject level and we found that γ_2 was very similar to our model. Since γ_1 was not significant, the reduced model B was also appropriate for our data. In both cases, while coefficients differed in magnitude across the models, the overall inference is the same.

We also considered reduced models that contain only one random effect. These models, C and D, are also included in the table. Reduced model C was only linked via the cluster level. The model is as follows,

$$y_{ijk} = Z_{ijk}^T \beta + a_i + e_{ijk}$$
$$logit(p_{ij}) = W_{ij}\alpha + \gamma_1 a_i.$$

Reduced model D was only linked via the subject level, i.e.

$$y_{ijk} = Z_{ijk}^T \beta + b_{ij} + e_{ijk}$$
$$logit(p_{ij}) = W_{ij}\alpha + \gamma_2 b_{ij}.$$

We found that the variance components, σ_a and σ_e, in reduced model C, and σ_b in reduced model D were inflated. Also, there were some differences in estimates between the two types of reduced models. For example, black patients have lower MAMC for models C and D which contained only one random effect. Reduced model C also had a lower MAMC time slope in magnitude when compared to the other models presented. These discrepancies showed that the correct modeling at the cluster level or the subject level is important. We also fitted models with the center level as a cluster level. Using unit level as a cluster level had a higher association even though the cluster level was not significant. We tried to fit a random slope in the time trend of MAMC, but it was not significantly different from zero.

For the goodness of fit test of the joint models, the performance of the proposed model was investigated further by computing the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The AIC and the BIC were applied to the HEMO data, which are shown in Table 5. The results show that our proposed model is better in terms of goodness of fit when compared with other reduced models.

Table 5: Results for HEMO data

	Our model			Red. model A			Red. model B			Red. model C			Red. model D		
Parameter	Est	SE	P	Est	SE	P	Est	SE	P	Est	SE	P	Est	SE	P
MAMC															
Intercept	23.844	0.523	0.000	23.814	0.518	0.000	23.767	0.525	0.000	23.762	0.494	0.000	23.871	0.540	0.000
Male	1.443	0.165	0.000	1.448	0.165	0.000	1.417	0.159	0.000	1.476	0.191	0.000	1.435	0.163	0.000
Back	0.803	0.196	0.000	0.810	0.195	0.000	0.805	0.199	0.000	0.740	0.204	0.001	0.789	0.212	0.001
High Kt/V	0.025	0.153	0.873	0.027	0.154	0.861	0.011	0.152	0.940	-0.067	0.162	0.681	0.004	0.151	0.980
High flux	0.130	0.129	0.318	0.130	0.129	0.316	0.121	0.127	0.343	0.113	0.133	0.401	0.118	0.129	0.367
ICED=3	-0.840	0.189	0.000	-0.838	0.189	0.000	-0.830	0.192	0.000	-0.744	0.211	0.001	-0.827	0.186	0.000
ICED=2	-0.413	0.205	0.049	-0.411	0.205	0.050	-0.397	0.201	0.054	-0.425	0.189	0.028	-0.407	0.203	0.050
Diabetes	0.907	0.210	0.000	0.904	0.211	0.000	0.907	0.209	0.000	0.862	0.207	0.000	0.922	0.217	0.000
Age	-0.002	0.008	0.789	-0.002	0.007	0.840	-0.002	0.007	0.838	-0.002	0.007	0.817	-0.003	0.007	0.635
Time	-0.122	0.031	0.000	-0.115	0.031	0.001	-0.123	0.031	0.000	-0.014	0.031	0.650	-0.123	0.031	0.000
Logit															
Intercept	-4.001	0.314	0.000	-3.923	0.303	0.000	-3.859	0.324	0.000	-3.801	0.314	0.000	-3.879	0.324	0.000
Male	0.230	0.136	0.096	0.238	0.128	0.070	0.232	0.122	0.064	0.241	0.119	0.047	0.233	0.123	0.064
Back	-0.144	0.152	0.350	-0.142	0.148	0.345	-0.117	0.142	0.411	-0.106	0.138	0.444	-0.118	0.143	0.410
High Kt/V	-0.001	0.103	0.994	0.003	0.101	0.976	0.031	0.104	0.776	0.025	0.101	0.808	0.030	0.104	0.772
High flux	-0.074	0.094	0.435	-0.072	0.089	0.423	-0.055	0.087	0.532	-0.060	0.084	0.481	-0.055	0.088	0.535
ICED=3	1.173	0.153	0.000	1.134	0.151	0.000	1.102	0.144	0.000	1.088	0.142	0.000	1.109	0.143	0.000
ICED=2	0.717	0.175	0.000	0.696	0.167	0.000	0.679	0.167	0.000	0.664	0.161	0.000	0.682	0.168	0.000
Diabetes	0.317	0.122	0.012	0.303	0.118	0.013	0.282	0.119	0.022	0.282	0.117	0.020	0.284	0.120	0.022
Age	0.052	0.005	0.000	0.051	0.005	0.000	0.051	0.005	0.000	0.050	0.005	0.000	0.051	0.005	0.000
γ_1	-1.178	1.125	0.300	-1.094	0.685	0.116				-0.209	0.193	0.285			
γ_2	-0.120	0.020	0.000				-0.124	0.019	0.000				-0.124	0.019	0.000
Var. Comp.															
σ_a	0.420	0.214	0.050	0.432	0.144	0.004	0.558	0.110	0.000	0.804	0.072	0.000			
σ_b	3.179	0.096	0.000	3.174	0.094	0.000	3.151	0.094	0.000				3.200	0.093	0.000
σ_e	1.832	0.115	0.000	1.832	0.115	0.000	1.832	0.115	0.000	3.500	0.096	0.000	1.832	0.115	0.000
AIC		27811.9			27850.8			27828.1			31038.4			27839.0	
BIC		27943.8			27977.2			27954.5			31159.3			27959.9	

2.5 DISCUSSION

Hierarchical structure is common in many biomedical studies, and ignoring the multi-level correlation dependence will lead to incorrect results. In this paper we developed a multi-level joint model of a longitudinal and a binary outcome. The two sub-models were linked with both the cluster and the subject level random effects. The results were compared with models that assumed only one level of dependence. The Gaussian quadrature technique was implemented using the aML software. Our simulation results showed that ignoring the correlations between outcomes can cause biased estimates.

In our motivating example of the HEMO study, the association between the repeated measurement of MAMC and the mortality rate were significant, but not for the association between unit level and mortality rate. In this example, the reduced model A was not appropriate because the model was linked only with the cluster level. Although the association of the cluster level was not significant in the HEMO study, it is still important to check for associations. We fitted the model with the center level as a cluster level. Using unit level as a cluster level had higher association even though it was not significant. We also tried to fit both a random intercept and a random slope in the time trend of MAMC, but random slope was not significantly different from zero. Our model assumed the cluster level random effect and subject level random effect were independent. For further investigation, the generalization of them could be considered.

3.0 INDIVIDUAL PREDICTIONS OF JOINT MODEL

3.1 INTRODUCTION

In joint modeling studies, investigators often want to predict an individual mortality rate in situations where future individualized treatments may be considered. There are some papers which present the subject-specific predictions in the joint modeling of longitudinal and survival outcomes. Taylor *et al.* [34] and Yu *et al.* [35] focused on individualized predictions for disease progression. They predicted future prostate-specific antigen (PSA) biomarkers and the predicted probability of cancer recurrence for censored and alive patients. The Markov chain Monte Carlo (MCMC) method has been applied to have individual draws. Garre *et al.* [36] proposed a joint latent class model for longitudinal and survival data with two latent classes and their predictions were better than other joint models. Proust-Lima and Taylor [37] focused on the dynamic prognostic tool from a joint latent class model and evaluated the predictive accuracy measures. Rizopoulos [38] assessed the predictive ability of the longitudinal marker of the joint model. Also, Horrocks *et al.* [27] considered the prediction of pregnancy in the joint model of longitudinal and binary outcomes.

In this chapter, we will fit our proposed model in WinBUGS, which is for Bayesian analysis using Markov chain Monte Carlo (MCMC) methods. Guo and Carlin [13] developed the method of Henderson into a fully Bayesian version using the Markov chain Monte Carlo technique. We will compare the estimates from using the Gaussian quadrature technique implemented in aML with the fully Bayesian approach via Markov Chain Monte Carlo (MCMC) methods.

Next, we focus on individualized predictions of mortality for a patient. The motivation for this work arose from individual prediction of our proposed model. Here, the longitudinal

measures of mid-arm muscle circumference (MAMC) serve as an indicator of health. By observing the longitudinal MAMC, we will be able to calculate the predicted probability of individual mortality.

3.2 METHODS

Let Y_{ijk} denote the k-th repeated measure for the j-th subject within the i-th cluster ($i = 1, 2, ..., n, j = 1, 2, ..., n_i,$ and $k = 1, 2, ..., m_{ij}$). Let p_{ij} denote the probability of response for the j-th subject of a binary zero-one outcome variable, X_{ij}. Let Z_{ijk} and W_{ij} denote the covariate vectors of fixed effects for the longitudinal outcome model and the logistic regression model, respectively. In this model, a_i and b_{ij} denote the random effects at the cluster and subject levels, respectively. They are assumed to be independent and identically distributed according to a normal distribution with mean 0 and corresponding variances, e.g. σ_a^2 and σ_b^2.

We define the joint model as

$$\begin{aligned} y_{ijk} &= Z_{ijk}^T\beta + a_i + b_{ij} + e_{ijk} \\ logit(p_{ij}) &= W_{ij}\alpha + \gamma_1 a_i + \gamma_2 b_{ij}, \end{aligned} \tag{3.1}$$

where β and α are unknown vectors of parameters. In this model γ_1 and γ_2 represent the association between the two models at each cluster and subject level. The error term, e_{ijk} is assumed to be N(0, σ_e^2) and independent of (a_i, b_{ij}, X_{ij}). Let O_{ij} denote the observed data for the i-th subject within the j-th cluster.

The individual patient has a different predicted probability, standard error and confidence bands using the Bayesian posterior distribution. The predicted probability of death for a subject is

$$\hat{P_{ij}} = \frac{exp(W_{ij}\hat{\alpha} + \hat{\gamma_1}\hat{a_i} + \hat{\gamma_2}\hat{b_{ij}})}{1 + exp(W_{ij}\hat{\alpha} + \hat{\gamma_1}\hat{a_i} + \hat{\gamma_2}\hat{b_{ij}})}, \tag{3.2}$$

where $\hat{\alpha}$ is the estimated coefficient of the logistic regression submodel, and $\hat{\gamma_1}$and $\hat{\gamma_2}$ are the estimated associations of the joint model at the cluster level and the subject level, respectively. The predicted subject level and cluster level random intercepts are $\hat{a_i}$ and $\hat{b_{ij}}$, respectively.

We will fit our proposed model in WinBUGS 1.4.3 which is based on the Markov Chain Monte Carlo (MCMC) method. It can be called from R using the R package 'R2WinBUGS'. WinBUGS has several advantages when fitting Bayesian models providing great flexibility, especially for multilevel modelling. There is also no restriction on number of levels of random effects. On the other hand, the program uses more computational time than a standard joint model fit using PROC NLMIXED in SAS. While this is not necessarily a fair comparison, since more complex models can not be fit using NLMIXED in SAS, it is certainly a consideration. Part of this computational burden arises from the fact that the estimates are simulated from posterior distributions. We used the estimates from the likelihood based method using the Gaussian quadrature techniques for the initial values of the parameters for the WinBUGs program to make the simulation computationally efficient.

Following the approach used by Guo and Carlin [13], we used proper but vague prior distributions, so that the priors will have minimal impact relative to data. For the longitudinal submodel, we assumed multivariate normal and inverse gamma priors for the main effects and the error variance, respectively. For the logistic regression submodel, normal priors were used. For the cluster level and the subject level random effects, the normal priors were used. Again, the normal priors were used for the associations of the joint model. To determine the accuracy of the predicted values, we computed ROC curves.

3.3 SIMULATION STUDY

We conducted a simulation study to better understand the performance of the proposed model when it is fit in the BUGS (Bayesian inference Using Gibbs Sampling) software using Markov chain Monte Carlo (MCMC) methods. The development of the BUGS project is currently focused on OpenBUGS, while WinBUGS is stable and not undergoing further

development. BRugs is an R package, that allows OpenBUGS to interface with R. Here, we used OpenBUGS 3.2.2 for the simulation study.

For all simulations we simulated data on 500 subjects with a cluster size of 50, so that there were 10 subjects within each of the clusters. The results presented are based on 600 simulated samples for each scenario. The model below was used for generating data:

$$y_{ijk} = \beta_0 + \beta_1 Z_{ij} + \beta_2 time + a_i + b_{ij} + e_{ijk}$$
$$logit(p_{ij}) = \alpha_0 + \alpha_1 Z_{ij} + \gamma_1 a_i + \gamma_2 b_{ij}.$$

Let y_{ijk} represent the repeated measurement at the integer time k, with k ranging from 1 to 5. We generated the subject-level covariate, Z_{ij}, from a binary distribution with probability 0.5. The coefficient parameters were as follows; $\beta = (\beta_0, \beta_1, \beta_2)^T = (-1, -.5, -.2)^T$. The random intercepts a_i and b_{ij} were also included in the model. We assumed $a_i \overset{\text{iid}}{\sim} N(0, \sigma_a^2)$ with $\sigma_a^2 = 1$ and $b_{ij} \overset{\text{iid}}{\sim} N(0, \sigma_b^2)$ with $\sigma_b^2 = 1$. Since the repeated measures from a subject j share the common random effects b_{ij} and the subjects from a cluster share the common random effects a_i, the correlation is induced from the random effects. Here, the covariance structure is compound symmetry where the variance at all time points is the same and the correlation between any two distinct measurements is the same.

The error term was $e_{ijk} \overset{\text{iid}}{\sim} N(0, \sigma_e^2)$ with $\sigma_e^2 = 1$. The logistic regression model included an intercept α_0 and the subject-level covariate Z_{ij} with coefficient $\alpha = (\alpha_0, \alpha_1)^T = (1.5, -1)^T$. Also, the random effects a_i, at the cluster level, and b_{ij}, at the subject level, were included in the model. We assumed that the coefficient parameters were as follows; $\gamma = (\gamma_1, \gamma_2)^T = (1, 1)^T$. We used two MCMC sampling chains of 10,000 iterations each, following a 5,000-iteration burn-in period.

Table 6 presents the simulation results. The estimates of our proposed model have very small bias and the coverage probability shows that the performance of our model is reasonable using a Bayesian approach. Figure 1 presents boxplots of simulation results with 600 samples providing graphs of the distribution of the estimates. The estimates of both the longitudinal model margin and the logistic regression model margin are unbiased, but the variance component of cluster level, σ_a, is highly variable with some outliers when compared to the other σ's. Also, the association parameters of the two models, γ_1 and γ_2, appear to

have larger variability than the other parameters. The area under the curve (AUC), which estimates the prediction performance, is 0.8326 using the average of 600 samples of the data.

For the simulation study we presented, we also note that the estimates using MCMC methods were unbiased with reasonable coverage probabilities. Based on the AUC results, the proposed model does a good job for prediction of binary outcome.

Table 6: Simulation Results

Parameter	Est	SD	CP Percent
$\beta_0 = -1$	-0.997	0.163	93.8
$\beta_1 = -0.5$	-0.507	0.103	94.3
$\beta_2 = 0.2$	0.199	0.014	96.5
$\alpha_0 = 1.5$	1.526	0.251	94.3
$\alpha_1 = -1$	-1.010	0.257	93.3
$\gamma_1 = 1$	1.037	0.152	95.5
$\gamma_2 = 1$	1.041	0.164	94.3
$\sigma_a = 1$	0.995	0.242	95.0
$\sigma_b = 1$	1.007	0.082	94.8
$\sigma_e = 1$	1.000	0.032	95.5

CP is the coverage probability of the 95 percent credible interval.

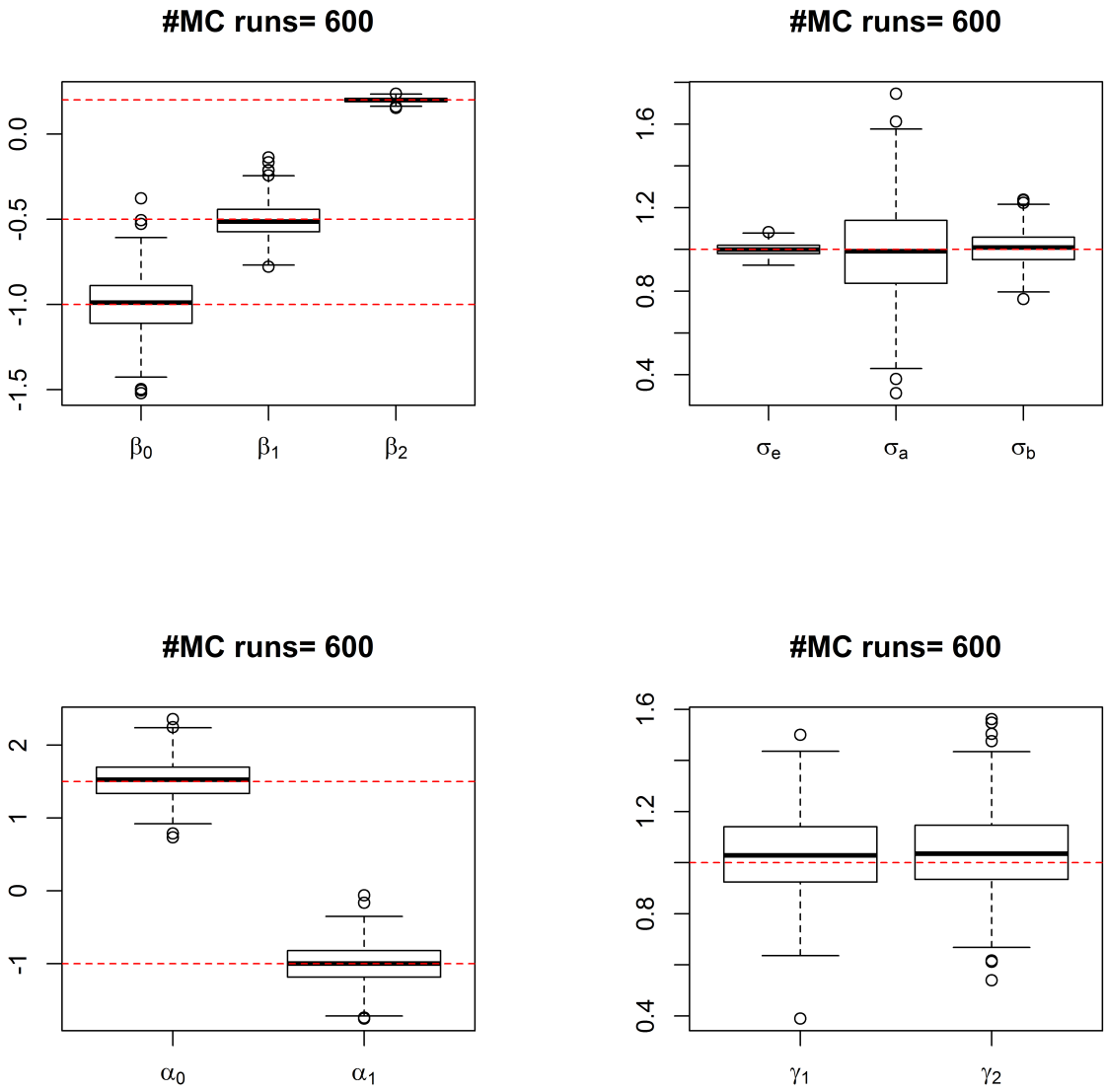

Figure 1: Boxplots of simulation results

3.4 APPLICATION

The HEMO study was a randomized controlled trial designed to identify the effects of dialysis dose and membrane flux on morbidity and mortality for patients undergoing chronic hemodialysis. The study enrolled 1846 patients nested within 75 units within 15 centers, randomized by dose (standard or high) and by membrane type (high or low). One question of secondary interest is the relationship between mid-arm muscle circumference and overall health. To address this we fit a joint model with the logistic outcome being mortality and the longitudinal component being mid-arm muscle circumference. A mixed model was fit to the data with the longitudinal measure of mid-arm muscle circumference (MAMC) as the outcome [33]. This measure was calculated using the following equation:

$$\begin{aligned}\text{MAMC (cm)} &= \text{mid-arm circumference (cm)} \\ &\text{- } 3.142 \times \text{triceps skinfold (TSF) (cm).}\end{aligned}$$

Our analysis included 1799 patients who had at least one measure of mid-arm muscle circumference (MAMC). Time is calculated from randomization date to the visit date for evaluation in years. The average number of follow-up visits is 3 (range 1-8) and the average follow-up time is 2.09 years (range: 0 - 6.36). Out of 1799 patients, 840 (46.7%) died during the study. The mean MAMC value is 24.6 (range 8.5 - 51.2).

The following baseline variables were included in both models: age at the first visit (mean 58 years), gender (44 % male), Index of Coexisting Disease (ICED) severity score which was calculated with diabetes excluded (36 % with a score of 1, 31 % with a score of 2 and 33 % with a score of 3), diabetic status (44 % diabetes) and race (63 % black). Also, the variables indicating treatment assignment, dose (standard vs. high dose) and membrane type (high or low flux membranes), were considered. Time in years since randomization date was also included in the longitudinal model.

Since patients receive similar treatment if they are in the same unit, there will be a within-cluster correlation for patients who are treated in the same unit (level 1). Also, the patients are nested within hospital (level 2) and finally the repeated measurements of mid-arm muscle circumference (MAMC) are nested within a patient (level 3). We assumed that

mortality might depend on the repeated measures of MAMC at both the unit level and the subject level. First, we built the longitudinal submodel with the outcome of MAMC and then we built the logistic regression submodel with the outcome of all-cause mortality. We linked the two sub-models using both subject and cluster level random effects. The results, shown in the right hand column of Table 7, are based on two MCMC sampling chains of 20,000 iterations each, following a 10,000-iteration burn-in period. The left hand column of Table 7 contains the results obtained from using gaussian quadrature techniques based on 50 quadrature points.

Most of the estimates are similar when comparing the gaussian quadrature techniques with the MCMC method. If we look at the results from the analysis based on the MCMC method, there was a significant linear decreasing pattern in MAMC over time. It decreased by 0.122 per year, since the 95% credible interval does not include 0. Patients who had a higher index of coexistent diseases (ICED) score had significantly higher mortality. Also, patients who had diabetes had higher mortality when compared to the patients with no diabetes by 0.320. We observed that patients with an older age had a higher mortality by 0.052.

We found that the random effects at both the cluster and the subject levels were significant for MAMC. The estimate of σ_a (cluster level) was 0.355 and σ_b (subject level) was 3.198. The longitudinal MAMC was negatively correlated with mortality rate at the subject level (γ_2) and at the unit level (γ_1). The association at the cluster level, γ_1, was not significant in the gaussian quadrature method. This is the only difference when we compare with the results from the two methods.

Figure 2 shows the individual trajectory of the MAMC for two selected patients, A and B, who are both alive. Patient A is 38 years old, black, male with an ICED score of 2 and no diabetes. Patient B is 68 years old, non black, male with an ICED score of 3 with diabetes. In this figure we see a clear difference in the trajectories of these 2 subjects. Patient A exhibited a rather stable pattern over time. The predicted probability of death is 0.2078 with a credible interval of (0.1247, 0.3117). Patient B, on the other hand, has a trajectory with a decreasing pattern at the end of measurement times. The predicted probability of death for this patient is 0.8405 with a credible interval of (0.7576, 0.9050). Note that patient

B has a much higher probability of death when compared with patient A. Figure 3 shows the ROC curve for the HEMO data. The area under the curve (AUC), estimates the prediction performance which is 0.7869 in our proposed model. Based on this result, the proposed model does a good job of predicting death in this cohort.

Table 7: Results for HEMO data: Gaussian Quadrature vs. MCMC

	Gaussian Quadrature		**MCMC**	
Parameter	Est.	P-value	Posterior Mean	95% CI
		MAMC		
Intercept	23.844	0.000	23.800	(23.010, 24.610)
Male	1.443	0.000	1.454	(1.117, 1.774)
Back	0.803	0.000	0.814	(0.472, 1.160)
High Kt/V	0.025	0.873	0.033	(-0.277, 0.343)
High flux	0.130	0.318	0.136	(-0.173, 0.445)
ICED=3	-0.840	0.000	-0.828	(-1.229, -0.425)
ICED=2	-0.413	0.049	-0.406	(-0.809, 0.001)
Diabetes	0.907	0.000	0.913	(0.577, 1.246)
Age	-0.002	0.789	-0.002	(-0.014, 0.009)
Time	-0.122	0.000	-0.122	(-0.160, -0.085)
		Logit		
Intercept	-4.001	0.000	-4.047	(-4.741, -3.436)
Male	0.230	0.096	0.229	(0.001, 0.461)
Back	-0.144	0.350	-0.150	(-0.412, 0.101)
High Kt/V	-0.001	0.994	-0.007	(-0.225, 0.208)
High flux	-0.074	0.435	-0.076	(-0.297, 0.137)
ICED=3	1.173	0.000	1.186	(0.910, 1.470)
ICED=2	0.717	0.000	0.728	(0.457, 1.003)
Diabetes	0.317	0.012	0.320	(0.094, 0.549)
Age	0.052	0.000	0.052	(0.044, 0.062)
γ_1	-1.178	0.300	-1.910	(-3.773, -0.759)
γ_2	-0.120	0.000	-0.120	(-0.159, -0.082)
		Var. Comp.		
σ_a	0.420	0.050	0.355	(0.177, 0.572)
σ_b	3.179	0.000	3.198	(3.074, 3.324)
σ_e	1.832	0.000	1.832	(1.790, 1.875)

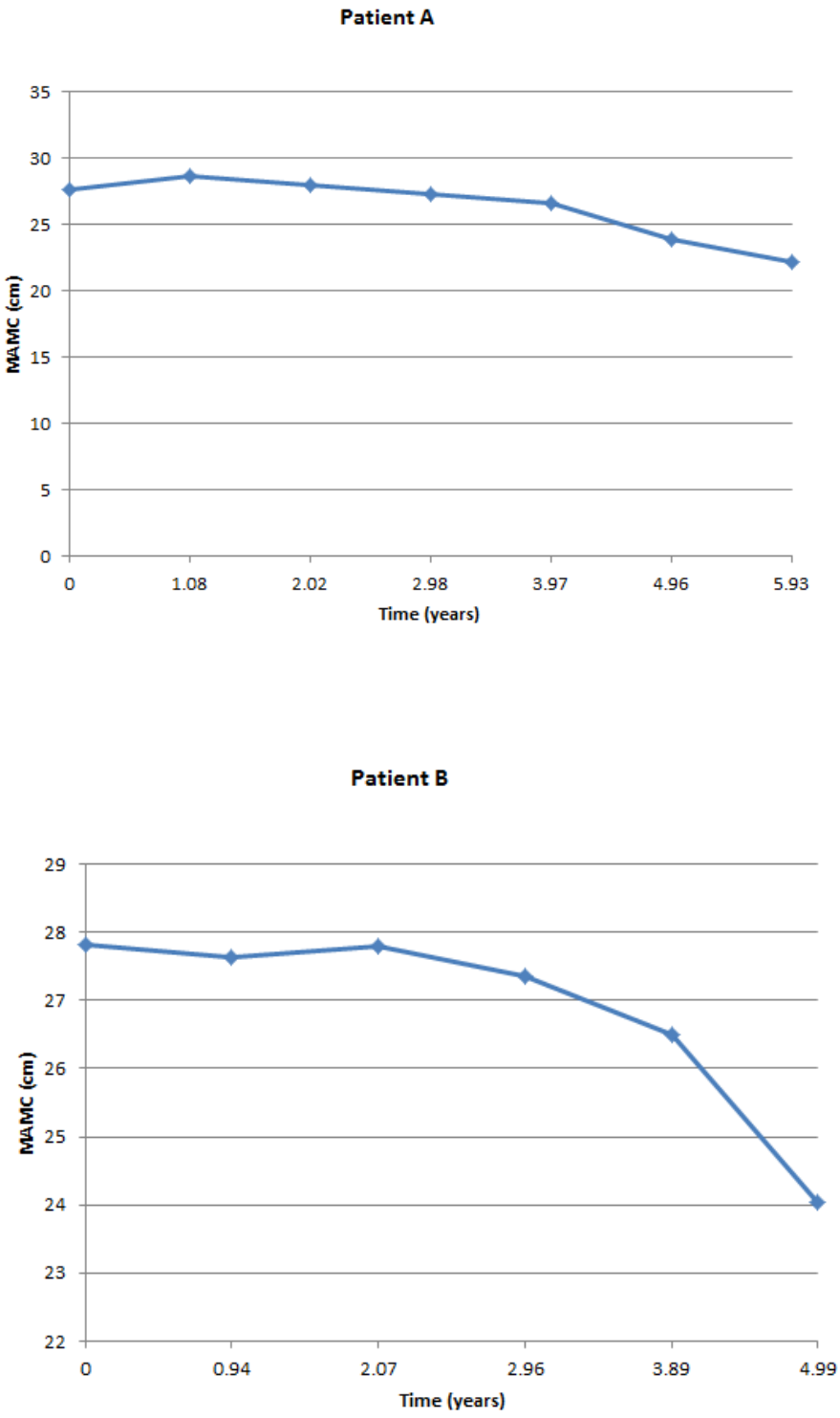

Figure 2: Observed MAMC for Patients A and B

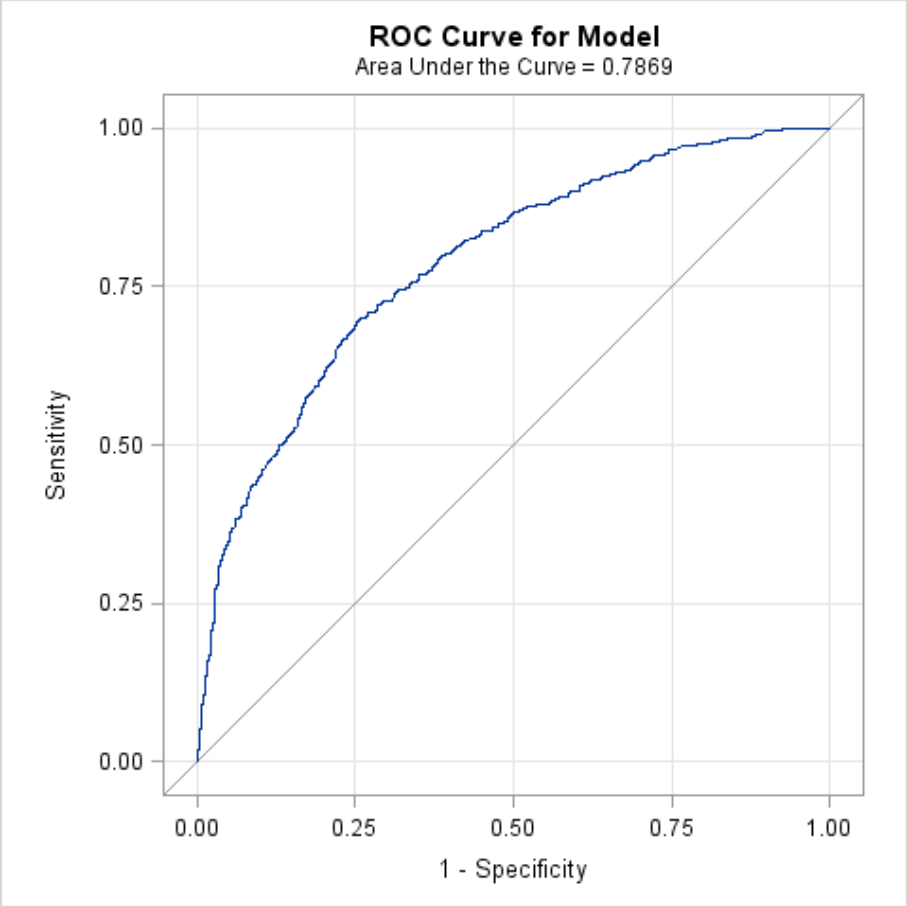

Figure 3: ROC Curve for HEMO data

3.5 DISCUSSION

The individualized predictions based on joint models are of increasing interest to many scientific investigators and the methods presented here provide an approach to address prediction. We focused on individual predictions estimated from the joint model of longitudinal and binary outcomes using the Bayesian approach available in the software package WinBUGS. This approach relies on Markov Chain Monte Carlo (MCMC) methods for the joint model analysis. WinBUGS provides great flexibility and there is no restriction on the number of random effects that can be included in the model. It can also fit more complex models, such

as multilevel models. We used proper but vague prior distributions in WinBUGS, so that the priors will have minimal impact relative to the data.

We focused on the individual predictions of our proposed model of multilevel joint model of longitudinal and binary outcomes. The individual prediction of mortality can be calculated from the draws in the Markov chain. We applied this method to the HEMO data to assess the relationship between the longitudinal trajectory of mid-arm muscle circumference (MAMC) and the predicted mortality. We observed a steep decreasing pattern of MAMC which may be an indicator of an increased mortality rate and suggest directions for the future treatment.

4.0 DISCUSSION

Hierarchical structure is common in many biomedical studies, such as longitudinal measures nested within subject and then nested within hospital. If we ignore the multi-level correlation dependence, it will lead to incorrect results. Here, we proposed a multi-level joint model of a longitudinal and a binary outcome. The two sub-models were linked with both the cluster and the subject level random effects. The results were compared with models that assumed only one level of dependence. The Gaussian quadrature technique was implemented using the aML software. The simulation study presented showed that the estimates were sensitive to the violation of the assumptions of the dependence structure between the longitudinal outcome and the binary outcome. We applied our model to the HEMO data which was a randomized controlled trial designed to identify the effects of dialysis dose and membrane flux on morbidity and mortality for patients undergoing chronic hemodialysis. The patients are nested within a unit and finally the repeated measurements of mid-arm muscle circumference (MAMC) are nested within a patient.

We also extended our approach to obtain individual predictions based on the proposed joint model. The motivation for this work arose from individual prediction of our proposed model. We fit our proposed model in software package (WinBUGS) for Bayesian analysis using Markov chain Monte Carlo (MCMC) methods. We focused on the individual predictions of our proposed model of the multilevel joint model of longitudinal and binary outcomes. The simulation study presented showed that based on the area under the curve (AUC) results, the proposed model is able to predict well for the binary outcome. The area under the curve (AUC) for the HEMO data also indicate that the model does a good job of predicting death in this cohort. Thus, the proposed method provides a mechanism for understanding the relationship between a longitudinal measure and a given binary outcome.

WinBUGS provides great flexibility and there is no restriction on the number of random effects that can be included in the model. It can also fit more complex models, such as multilevel models. On the other hand, the program uses more computational time than a standard joint model fit using PROC NLMIXED in SAS. While this is not necessarily a fair comparison, since more complex models can not be fit using NLMIXED in SAS, it is certainly a consideration.

BIBLIOGRAPHY

[1] Neuhausy A, Augustin T, Heumann C, Daumer M. A Review on Joint Models in Biometrical Research. *University of Munich, Discussion Paper* 2006; **506**:SFB 386.

[2] Tsiatis AA, DeGruttola V, Wulfsohn MS. 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* 1995; **90**:27–37.

[3] Schluchter MD. Methods for the Analysis of Informatively Censored Longitudinal Data. *Statistics in Medicine* 1992; **11**:1861–1870.

[4] Dempster AP, Laird NM, Rubin DB. Maximum Likelihood from Incomplete Data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* 1977; **39**:1–38.

[5] DeGruttola V, Tu XM. Modelling Progression of CD4-Lymphocyte Count and Its Relationship to Survival Time. *Biometrics* 1994; **50**:1003–1014.

[6] Wulfsohn MS, Tsiatis AA. A Joint Model for Survival and Longitudinal Data Measured with Error. *Biometrics* 1997; **53**:330–339.

[7] Henderson R, Diggle P, Dobson A. Joint Modelling of Longitudinal measurements and event time data. *Biostatistics* 2000; **1**:465–480.

[8] Lin H, McCulloch CE, Mayne ST. Maximum Likelihood Estimation in the Joint Analysis of Time-to-Event and Multiple Longitudinal Variables. *Statistics in Medicine* 2002; **21**:2369–2382.

[9] Song X, Davidian M, Tsiatis AA. A Semiparametric Likelihood Approach to Joint Modeling of Longitudinal and Time-to-Event Data. *Biometrics* 2002; **58**:742–753.

[10] Hsieh F, Tseng YK, Wang JL. Joint Modeling of Survival and Longitudinal Data: Likelihood Approach Revisited. *Biometrics* 2006; **62**:1037–1043.

[11] Faucett CL, Thomas DC. Simultaneously Modelling Censored Survival Data and Repeatedly Measured Covariates: a Gibbs Sampling Approach. *Statistics in Medicine* 1996; **15**:1663–1685.

[12] Xu J, Zeger SL. Joint Analysis of Longitudinal Data Comprising Repeated Measures and Times to Events. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 2001; **50**:375–387.

[13] Guo X, Carlin BP. Separate and Joint Modeling of Longitudinal and Event Time Data Using Standard Computer Packages. *The American Statistician* 2004; **58**:16–24.

[14] Hogan JW, Laird NM. Model-based Approaches to Analysing Incomplete Longitudinal and Failure Time Data. *Statistics in Medicine* 1997; **16**:259–272.

[15] Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: John Wiley, 2002.

[16] Tsiatis AA, Davidian M. A Semiparametric Estimator for Proportional Hazards Model with Longitudinal Covariates Measured with Error. *Biometrika* 2001; **88**:447–458.

[17] Tsiatis AA, Davidian M. Joint Modelling of Longitudinal and Time-to-Event Data: an Overview. *Statistica Sinica* 2004; **14**:809–834.

[18] Stefanski LA, Carroll RJ. Conditional Scores and Optimal Scores for Generalized Linear Measurement- Error Models. *Biometrika* 1987; **74**:703–716.

[19] Song X, Davidian M, Tsiatis AA. An Estimator for the Proportional Hazards Model with Multiple Longitudinal Covariates Measured with Error. *Biometrics* 2002; **3**:511–528.

[20] Littell R, Milliken G, Stroup W, Wolfinger R, Schabenberger O. *SAS for Mixed Models*. SAS Institute Inc., 2006.

[21] Vonesh EF, Greene T, Schluchter MD. Shared Parameter Models for the Joint Analysis of Longitudinal Data and Event Times. *Statistics in Medicine* 2006; **25**:143–163.

[22] Liu L. Joint Modeling Longitudinal Semi-Continuous Data and Survival, with Application to Longitudinal Medical Cost Data. *Statistics in Medicine* 2009; **28**:972–986.

[23] Liu L, Huang X. The use of Gaussian Quadrature for Estimation in Frailty Proportional Hazards Models. *Statistics in Medicine* 2008; **27**:2665–2683.

[24] Wang CY, Wang N, Wang S. Regression Analysis When Covariates Are Regression Parameters of a Random Effects Model for Observed Longitudinal Measurements. *Biometrics* 2000; **56**:487–495.

[25] Li E, Zhang D, Davidian M. Conditional Estimation for Generalized Linear Models When Covariates Are Subject-Specific Parameters in a Mixed Model for Longitudinal Measurements. *Biometrics* 2004; **60**:1–7.

[26] Li E, Wang N, Wang NY. Joint Models for a Primary Endpoint and Multiple Longitudinal Covariate Processes. *Biometrics* 2007; **63**:1068–1078.

[27] Horrocks J, van Den Heuvel MJ. Prediction of Pregnancy: A Joint Model for Longitudinal and Binary Data. *Bayesian Analysis* 2009; **4**:523–538.

[28] Wannemuehler KA, Lyles RH, Manatunga AK, Terrell ML, Marcus M. Likelihood-Based Methods for Estimating the Association between a Health Outcome and Left- or Interval-Censored Longitudinal Exposure Data. *Statistics in Medicine* 2009; **29**:1661–1672.

[29] Ratcliffe SJ, Guo W, Have TRT. Joint Modeling of Longitudinal and Survival Data via a Common Frailty. *Biometrics* 2004; **60**:892–899.

[30] Liu L, Ma JZ, OQuigley J. Joint Analysis of Multi-Level Repeated Measures Data and Survival: an Application to the End Stage Renal Disease (ESRD) Data. *Statistics in Medicine* 2008; **27**:5679–5691.

[31] Lillard L, Panis C. *aML Multilevel Multiprocess Statistical Software, Version 2.0*. Econ Ware: Los Angeles, CA, 2003.

[32] Eknoyan G, Beck GJ, Cheung AK, Daugirdas JT, Greene T, Kusek JW, Allon M, Bailey J, Delmez JA, Depner TA, *et al.*. Effect of Dialysis Dose and Membrane Flux in Maintenance Hemodialysis. *The New England Journal of Medicine* 2002; **347**:2010–2019.

[33] Noori N, Kopple JD, Kovesdy CP, Feroze U, amd Sameer B Murali JJS, Luna A, Gomez M, Luna C, Bross R, Nissenson AR, *et al.*. Mid-Arm Muscle Circumference and Quality of Life and Survival in Maintenance Hemodialysis Patients. *Clinical Journal of the American Society of Nephrology* 2010; **5**:2258–2268.

[34] Taylor JM, Yu M, Sandler HM. Individualized Predictions of Disease Progression Following Radiation Therapy for Prostate Cancer. *Journal of Clinical Oncology* 2005; **23**:816–825.

[35] Yu M, Taylor JM, Sandler HM. Individual Prediction in Prostate Cancer Studies Using a Joint Longitudinal Survival-Cure Model. *Journal of the American Statistical Association* 2008; **103**:178–187.

[36] Garre FG, Zwinderman AH, Geskus RB, Sijpkens YWJ. A joint Latent Class Changepoint Model to Improve the Prediction of Time to Graft Failure. *Journal of the Royal Statistical Society, Series A* 2008; **171**:299–308.

[37] Proust-Lima C, Taylor JM. Development and Validation of a Dynamic Prognostic Tool for Prostate Cancer Recurrence Uing Repeated Measures of Posttreament PSA: a Joint Modeling Approach. *Biostatistics* 2009; **10**:535–549.

[38] Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics* 2011; **67**:819–829.