THE APPLICATION OF JOINT LONGITUDINAL-SURVIVAL MODELS TO BIVARIATE DECISION MAKING IN CLINICAL TRIALS

by

ELISE GUNZBURGER

B.S., University of Florida, 2011

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Biostatistics Program

2015

This thesis for the Master of Science degree by

Elise Gunzburger

has been approved for the

Biostatistics Program

by

John Kittelson, Chair

Elizabeth Juarez-Colunga

Edward Bedrick

Date 12/18/2015

Gunzburger, Elise (M.S., Biostatistics)

The Application of Joint Longitudinal-Survival Models to Bivariate Decision Making in Clinical Trials

Thesis directed by Professor John Kittelson

ABSTRACT

Although multiple endpoints are typically considered in the design and interpretation of clinical trials, decision criteria are usually based on a single primary endpoint. This thesis aims to examine methods for constructing decision criteria using a bivariate endpoint so that decisions can consider joint effects on both the primary and key secondary endpoints. Specifically, decision criteria for the joint effects of a treatment on a survival endpoint and a companion longitudinal endpoint are considered. Bivariate decision criteria are explored based on three different methods for linking longitudinal and survival submodels using readily available software. The methods are illustrated with the application to a well know AIDs dataset as well as to a recent major cardiovascular trial.

The form and content of this abstract are approved. I recommend its publication.

Approved: John Kittelson

TABLE OF CONTENTS

LIST OF TABLES

TABLE

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Multiple endpoints are typically considered in the design and analysis of clinical trials for various reasons. Motives for multiple endpoints include composite endpoints (e.g. benefit-risk tradeoff), when establishing mechanism and when there are key secondary endpoints of interest along with the primary endpoint. Decision criteria, including interim decision criteria, are commonly based on only one primary endpoint, although information on all specified endpoints are obtained in the trial. However, using a bivariate endpoint to construct decision criteria in clinical trial designs allows joint effects on both the primary and key secondary endpoints to be considered. This may be particularly useful in survival settings where information on the effect of treatment on the primary survival endpoint may accumulate slowly, whereas information on the effect of treatment on a key secondary endpoint accumulates more rapidly. Additionally, secondary endpoints in a survival setting are often companion longitudinal endpoints that are intermediary on the proposed mechanistic pathway or are supposed surrogate endpoints for the primary survival endpoint.

In the case where there is a primary endpoint and a key secondary endpoint, decision criteria can be based on a bivariate outcome space that is constructed using both endpoints. This bivariate outcome space takes into account treatment effects on both endpoints and can be divided into regions of benefit and harm for both endpoints. As seen in Figure 1, a bivariate outcome space can be constructed with a survival endpoint and a longitudinal endpoint.

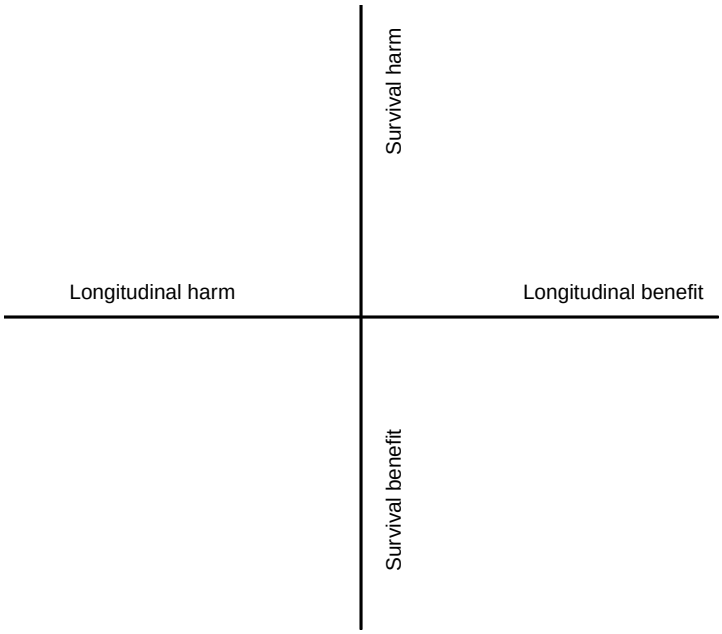


Figure 1. **Bivariate outcome space**: Example for survival and longitudinal endpoints

Hence, instead of basing decision criteria on the typical univariate null hypothesis, a bivariate null hypothesis can be constructed and used to make decisions.

One way in which to construct a bivariate null is using the typical mediation framework where the total effect of treatment, X, on the primary endpoint, Y, is mediated through the key secondary endpoint, Z, all for the ith subject

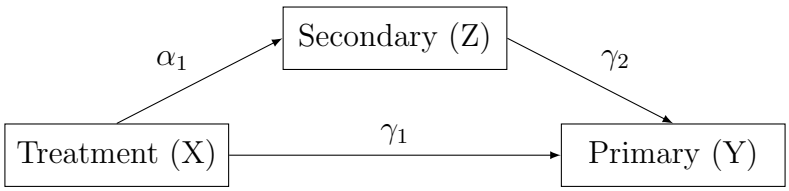
$$z_i = \alpha_0 + \alpha_1 x_i + \varepsilon_{1i}$$

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_{2i}$$

$$y_i = \gamma_0 + \gamma_1 x_i + \gamma_2 Z_i + \varepsilon_{3i}$$

where $i = 1, ..., n$ and α_0, α_1, etc. are the given parameters. This gives the division of

treatment effects on the primary endpoint into the direct effect, γ_1, indirect effect, γ_2, and total effect, $\beta_1 = \gamma_1 + \gamma_2\alpha_1$.



Using this mediation construction, a weighted composite bivariate null can be formed that ties together the treatment effects on both endpoints. Let $g(\beta_1, \alpha_1) = \gamma$ denote a level curve of a function that divides the bivariate outcome space into regions of benefit and lack of benefit. If $g(\beta_1, \alpha_1) = \beta_1 + \gamma_2\alpha_1$, then $g(\beta_1, \alpha_1) = \gamma_0$ is a linear null hypothesis (with usual choice $\gamma_0 = 0$, as seen in Figure 2) that corresponds to the division of total effects into direct and indirect effects.

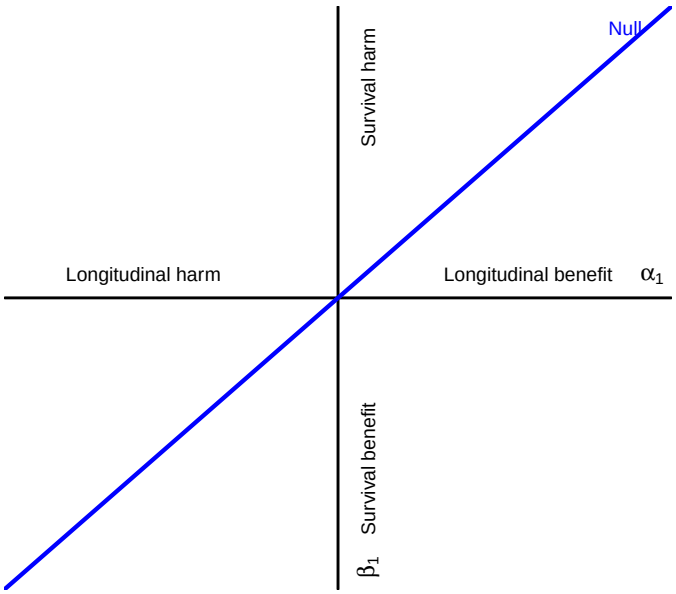


Figure 2. **Bivariate outcome space**: Example of a linear null hypothesis

However, a linear null implies an infinite tradeoff between the two endpoints, which is not clinically practical. Alternatively, a hyperbolic null avoids this deficit while still taking into account both endpoints and can be constructed by letting

$$g(\beta_1, \alpha_1) = (\frac{\beta_1 - \Delta_\beta}{\beta_\emptyset - \Delta_\beta})(\frac{\alpha_1 - \Delta_\alpha}{\alpha_\emptyset - \Delta_\alpha}) \quad .$$

One chooses $\Delta_\beta, \Delta_\alpha, \beta_\emptyset$, and $\alpha_\emptyset$ such that the null passes through the non-inferiority bound for β_1 when there is substantial benefit on α_1, through the non-inferiority bound for α_1 when there is substantial benefit on β_1 and through a specified $(\beta_\emptyset, \alpha_\emptyset)$ in order to obtain the desired curvature. Figure 3 shows an example of a hyperbolic null.

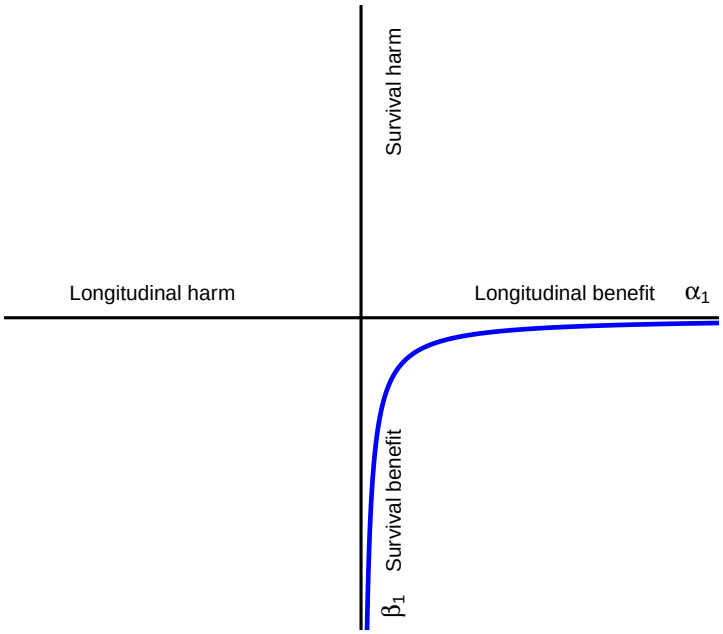


Figure 3. **Bivariate outcome space**: Example of hyperbolic null hypothesis

The hyperbolic null can be extended to construct hyperbolic interim boundaries in the

bivariate outcome space in order to make decisions at interim analyses based on both endpoints.

Based on a hyperbolic null, decisions concerning the results of a trial can be made using a confidence rectangle. A 95% confidence rectangle is constructed with the critical value, parameter estimates for α_1 and β_1 and their standard errors from the fixed sample or interim analysis results with the following construction:

- Null: $g(\beta_1, \alpha_1) = \gamma_0 \quad$ (with usual choice $\gamma_0 = 1$)
- Fixed sample where the critical value satisfies
  $g[\beta_1 - 1.96\ SE_F(\beta_1),\ \alpha_1 - 1.96\ SE_F(\alpha_1)] = \gamma_0$
- Group sequential where the critical value at the interim analysis satisfies:
  $g[\beta_1 - 1.96\ SE_{IA}(\beta_1),\ \alpha_1 - 1.96\ SE_{IA}(\alpha_1)] = \gamma_0$

where SE_F is the standard error for a fixed design and SE_{IA} is the standard error at the specified interim analysis. Additionally, the degree of early conservatism can be adjusted for the group sequential confidence rectangles. Although not explicitly stated, confidence rectangles are typically how results are reported by presenting the separate confidence intervals for each endpoint. However, confidence rectangles do not account for any correlation between the two endpoints and hence confidence ellipses may be used instead to account for this correlation. A 95% confidence ellipse can be constructed with the parameter estimates for α_1 and β_1 and their variance-covariance matrix with the following construction

- Centered at: (α_1, β_1)
- Length along α_1 axis: $2\sqrt{\lambda_{\alpha_1}\chi^2_{2,0.975}}$
- Length along β_1 axis: $2\sqrt{\lambda_{\beta_1}\chi^2_{2,0.975}}$
- Angle of rotation: $\arctan[\mathbf{v_1}(\beta_\mathbf{1})/\mathbf{v_1}(\alpha_\mathbf{1})]$

where λ_{α_1} and λ_{β_1} are the eigenvalues for the variance-covariance matrix and $\mathbf{v_1}$ is the eigenvector corresponding to the largest eigenvalue in magnitude. The same construction can be used to create 95% confidence ellipses at interim analyses.

Models that incorporate both survival and longitudinal endpoints into a single model may be beneficial compared to fitting separate models when it comes to making decisions, as information is included on both endpoints to provide a better estimate of treatment effect on survival. An extension of the Cox proportional-hazards model for a time-dependent covariate is an established way in which to incorporate a longitudinal endpoint into a survival model in order to estimate the adjusted treatment effect on survival. However, this time-dependent covariate model is only applicable to longitudinal responses that are external to the survival endpoint. The value of an external covariate, such as age or treatment group, is not affected by the occurrence of an event and can be known whether an individual is event free or not. In the setting of interest here, the longitudinal endpoint is internal, generated by a process directly related to the failure mechanism and measured with error. Hence it is not appropriate for use in a time-dependent Cox model.

This thesis aims to examine various ways of constructing joint models for a survival endpoint and companion internal longitudinal endpoint measured with error. Specifically, joint longitudinal-survival models from the random effects shared parameter models framework including current-value, slope dependent and random effects coefficients shared parameters are explored. Software to fit these joint models is available in the form of the JM package in R, the SAS Macro %JM and NLMIXED in SAS/STAT$^{\text{TM}}$ software. Motivation stems from large cardiovascular clinical trials where only one endpoint has been used to make decisions. In the dal-Outcomes trial, decision criteria on the effectiveness of treatment were based only on the primary survival endpoint, while the secondary endpoint, which could serve as a surrogate, was not incorporated (Schwartz et al., 2012). Similarly, in recent trials of PCSK9 inhibitors, decisions were made based solely on the surrogate endpoint without incorporating a survival endpoint. Including both endpoints

into a bivariate outcome space in these settings may have benefits, especially when information growth on the longitudinal endpoint is much more rapid than on the primary survival endpoint.

CHAPTER II

STATISTICAL METHODS

Longitudinal Submodel

Linear mixed effects models are commonly used to model continuous longitudinal data in order to account for repeated measurements made on the same individuals. The fixed effects part of the model represents the mean response, while the random effects part represents the individual level responses. Let $z_i(t)$ denote the observation of the longitudinal endpoint for the ith subject at time t, α^T denote the vector of coefficients for the fixed effects parameters and X_i denote the design matrix for the fixed effects. Then the linear mixed effect model for the longitudinal endpoint is

$$z_i(t) = \alpha^T X_i(t) + b_{i0} + b_{i1}(t) + \varepsilon_i(t) \qquad \varepsilon_i(t) \sim \text{iid } N(0, \sigma^2)$$

where b_{i0} represents the random effect intercept and $b_{i1}(t)$ the random effect slope for time, which are distributed as a mean-zero bivariate normal. Given the random effects, the longitudinal responses on an individual are assumed to be independent

$$p(z_i \mid b_i; \theta) = \prod_j p(z_i(t_{ij}) \mid b_i; \theta)$$

where θ is the vector of parameters. Additionally, conditional on b_{i0} and b_{i1}, the expected value of the longitudinal response for the ith subject at time t, $m_i(t)$, is given by

$$m_i(t) = E[z_i(t)] = \alpha^T X_i(t) + b_{i0} + b_{i1}(t) \ .$$

Survival Submodel

Let $h_i(t)$ represent the hazard of an event and $h_0(t)$ the baseline risk function. The general formulation of a proportional hazards time-to-event model is then

$$h_i(t) = h_0(t) \exp\{\gamma^T w_i\}$$

where γ^T denotes the vector of coefficients for the vector of baseline covariates w_i. An exponential regression model of the parametric proportional hazards family can be formulated by assuming the baseline risk function is constant over time, $h_0(t) = \lambda_0$, thus giving the model

$$h_i(t) = \lambda_0 \exp\{\gamma^T w_i\} \ .$$

The parametric exponential regression model is a relatively simple survival model and hence used in this thesis to examine joint models without the added complications of elaborate survival models.

Joint Longitudinal-Survival Models

Joint longitudinal-survival models can be formed where the association between the two endpoints is due to latent variables. In the formulation presented here, the longitudinal and survival processes are linked with a latent bivariate Gaussian process as the random effects b_i underlie both processes. Given the random effects, the longitudinal responses and the time-to-event responses are assumed to be independent, thus

$$p(T_i, \delta_i, z_i \ | b_i; \theta) = p(T_i, \delta_i \ | b_i; \theta) \ p(z_i \ | b_i; \theta)$$

where T_i is the observed event or censoring time and δ_i denotes the event indicator. Hence, the joint log-likelihood contribution of the ith subject is expressed as

$$\log\ p(T_i, \delta_i, z_i\ |\theta) = \log \int p(T_i, \delta_i\ |b_i; \theta)\ [\prod_j p(z_i(t_{ij})\ |b_i; \theta)]\ p(b_i; \theta) db_i\ \ .$$

Joint models of this type belong to the random effects shared parameter models framework as both submodels share the same random effects in their probability specification. However, there are numerous ways in which to link the longitudinal and survival submodels in the random effects shared parameter models framework. One way in which to link the submodels is through the current-value of the expected value of the longitudinal endpoint, $m_i(t)$, at a given time t. This is a current-value shared parameter model and is specified by

$$h_i(t) = h_0(t)\ \exp\{\gamma^T w_i + \eta\ m_i(t)\}$$

where η represents the association parameter for the expectation of the longitudinal endpoint and survival. Hence with this formulation, one assumes that the risk of an event at time t is dependent on the true value of the longitudinal endpoint at that time.

Another formulation of a random effects shared parameter model includes the linkage of the two submodels by the trajectory over time of the longitudinal endpoint. This model is the slope dependent shared parameter model and assumes that the risk of an event is dependent on the rate of increase or decrease (slope) of the longitudinal response. It is expressed by

$$h_i(t) = h_0(t)\ \exp\{\gamma^T w_i + \eta\ m_i'(t)\}$$

where $m_i'(t) = dm_i(t)/dt$, the first derivative of the expected longitudinal response with respect to time, and η is the estimate of the association between $m_i'(t)$ and survival.

Yet another way to link the submodels is with the isolated random effects b_i from the longitudinal linear mixed effects model. This model is the random effects coefficients model, specified by

$$h_i(t) = h_0(t) \exp\{\gamma^T w_i + \eta_1 \, b_{i0} + \eta_2 \, b_{i1}\}$$

where η_1 and η_2 estimate the association between the submodels induced by the random intercept and slope, respectively.

CHAPTER III

SOFTWARE

JM package for R

The JM package for R was developed for the use of joint modeling of longitudinal and survival data (Rizopoulos, 2010). Of the models discussed above, the package allows for the fitting of the current-value and slope dependent shared parameter models as well as a model that includes both the current-value and slope. The longitudinal endpoint can only be normal and a variance-components structure for the random effects variance-covariance structure is mandatory. Both relative risk and accelerated failure time formulations can be fit with an exponential, Weibull, piecewise-constant or spline-approximated baseline risk function specified. Additionally, the baseline risk function can be left unspecified to fit a Cox proportional-hazards model.

Estimates are determined by maximizing the log-likelihood of the joint distribution of the survival and longitudinal outcomes by approximating the integral with the Gauss-Hermite rule (default) or fully exponential Laplace approximation. Maximization of the log-likelihood is performed using the EM algorithm for a fixed number of iterations and then switching to a quasi-Newton algorithm until convergence is attained. Alternatively, one can choose to only use the EM algorithm for maximization if preferred. Initial values for parameters are taken from the estimates from the separate longitudinal and survival models. Using the `vcov` command in R on the joint model object, the variance-covariance matrix of the parameter estimates from the joint model can be obtained. For more information on this package, see the JM package documentation and reference paper (Rizopoulos, 2010; Rizopoulos, 2015).

%JM Macro for SAS

The %JM Macro for SAS software was developed to fit random effects shared parameter models for longitudinal and survival data using the NLMIXED procedures (Garcia-Hernandez and Rizopoulos, 2105). This Macro allows for the fitting of current-value, slope dependent and random effects coefficients models, among others. The longitudinal endpoint can be normal, binary, binomial or Poisson. Additionally, the variance-covariance structure of the random effects can have a variance-component structure, ante-dependence structure, or unstructured form. The relative risk formulation is used for the survival model and the baseline risk function must be specified, including exponential and Weibull options.

Similarly to the JM package for R, estimates are obtained by maximizing the log-likelihood using Gauss-Hermite quadrature to approximate to the integral by default but allowing other approximation methods such as the first-order method of Beal and Sheiner. The default method to maximizing the log-likelihood is the dual quasi-Newton optimization technique. Other optimization algorithms are available, including Newton-Raphson, but the EM algorithm is not an option. Initial values for parameters are taken from parameter estimates from the separate longitudinal and survival models. By specifying the ECOV option for NLMIXEDOptions macro parameter, the variance-covariance matrix for the parameter estimates of the joint model can be displayed.

As the %JM Macro and JM package were programed independently, analysis can be done with one software program while the other program is used as a corresponding quality control check to verify output. For more information on the JM% Macro, see the associated reference manual and website (Garcia-Hernandez and Rizopoulos, 2015).

NLMIXED in SAS

As an alternative to using the %JM Macro in SAS software, it is possible to fit some basic joint models directly with the NLMIXED procedures. Example SAS software code on joint modeling provides an example to fit a random effects coefficients shared parameter model (Guo and Carlin, 2004; Schabenberger, 2003). The marginal log-likelihood is estimated using Gaussian quadrature and the default optimization algorithm is the dual quasi-Newton technique. Initial values for parameters are assigned a default value of one; however, initial values could be specified if desired. More information can be found in the SAS software documentation for the NLMIXED procedure and example code (Schabenberger, 2003).

CHAPTER IV

RESULTS

Clinical Programs for Clinical Research on AIDS

The Clinical Programs for Clinical Research on AIDS (CPCRA) conducted a clinical trial to compare the efficacy and safety of two antiretroviral drugs. This dataset was used as an example dataset for analysis in the JM package and %JM Macro reference manuals, as well as in a paper on joint modeling in a Bayesian framework (Rizopoulos, 2010; Garcia-Hernandez and Rizopoulos, 2015; Guo and Carlin, 2004). Hence, the CPCRA data were used as a preliminary step to gain familiarity with and explore the different software methods as well as to do a basic comparison of their results. The current-value, slope dependent and random effects coefficients shared parameter models were each fit for this data using two different software programs.

The two antiretroviral study treatment drugs were didanosine (ddI) and zalcitabine (ddC). The study population was patients who were intolerant of or failed zidovudine therapy. There were 467 HIV-infected individuals who met entry requirements in the trial and were randomized to one of the two treatments. CD4 cell count is commonly used as a surrogate for AIDS progression where a count below a certain threshold is used as a biomarker of progression. CD4 cell count was recorded at randomization, 2, 6, 12 and 18 month visits. The primary outcome was death. At the end of the trial 188 (40.3%) patients had died.

For the CPCRA data, the linear mixed effects model used for CD4 count was

$$E[\sqrt{CD4_i(t)}] = \alpha_0 + \alpha_1 t + \alpha_2 t * tx_i + b_{i0} + b_{i1} t \qquad G = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

where t represents time and tx_i the indicator for treatment group for the ith subject where ddC was used as the reference group. Time was measured in months. A random intercept and slope for time were included in the model with a variance-components

structure for the random effects variance-covariance matrix. As CD4 counts are known to have right skewed distributions, the square root of CD4 count was used. For the survival submodel for the risk of death, a parametric exponential regression model was defined with constant baseline risk, giving the model

$$h_i(t) = \lambda_0 \exp\{\gamma_0 + \gamma_1 t x_i\} \quad .$$

Parameter estimates from fitting the linear mixed effects model and exponential regression model separately using `lme` and `survreg` in R 3.2.2, respectively, were given in Table 1.

Table 1. **CPCRA**: Separate longitudinal and exponential survival models parameter estimates (standard errors) with random effects standard deviations

	Separate models
Fixed Effects	
Intercept (α_0)	2.51 (0.043)*
Time (α_1)	-0.038 (0.0044)*
Time*Tx (α_2)	0.0082 (0.0063)
Random Effects	
Intercept (b_0)	0.87
Time (b_1)	0.036
Residual	0.37
Event Process(log h)	
Intercept (γ_0)	-3.55 (0.11)*
Tx (γ_1)	0.20 (0.15)
- Log Lik.	-1343, -835

*Significant at $p < 0.05$ level.

For the longitudinal model, ddI treatment benefit over ddC would be indicated by an increase in $\sqrt{CD4}$ count and ddI survival benefit indicated by a negative log(hazard), and therefore increased survival as compared to the ddC group. There was not a significant difference in the time trajectory of $\sqrt{CD4}$ count between ddI and ddC treatment groups. However, this estimate was in the direction of benefit of ddI over ddC. Treatment group was also not a significant predictor of survival [log(hazard)], but the estimate indicated

decreased survival in the ddI group compared to ddC. This was seen by the placement of the 95% confidence rectangle from the separate models in the bivariate outcome space in Figure 1 constructed from the 95% confidence intervals for the ddI treatment effect on $\sqrt{CD4}$, α_2, and treatment effect on the log(hazard), γ_1.

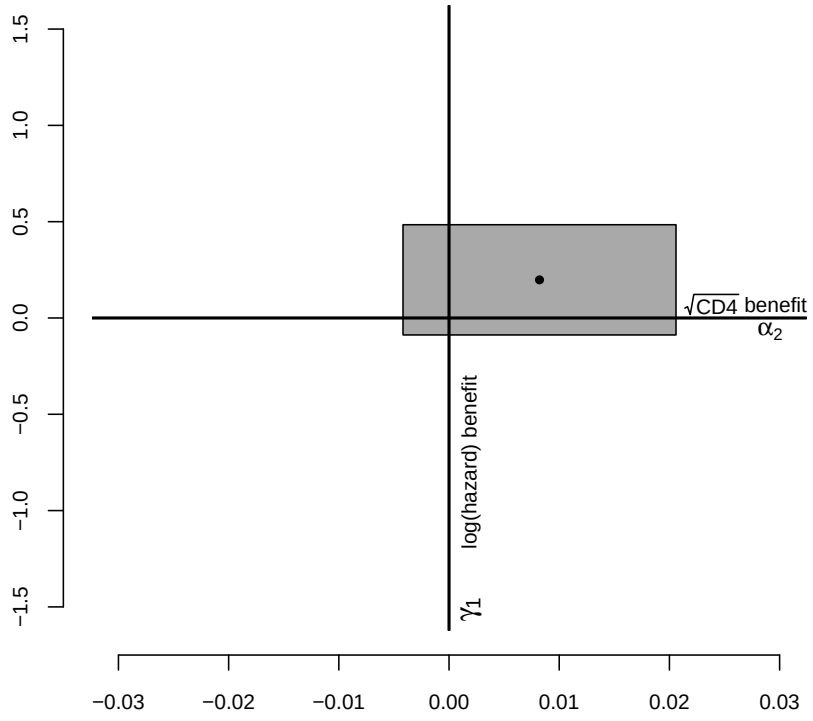


Figure 4. **CPCRA**: Separate longitudinal and survival models 95% confidence rectangle

First the joint longitudinal-survival model framework was explored with the CPCRA data by fitting the current-value shared parameter model as formulated by

$$h_i(t) = \lambda_0 \exp\{\gamma_0 + \gamma_1 tx_i + \eta\, m_i(t)\}$$

where $m_i(t)$ was the expected value of $\sqrt{CD4}$ count for the ith subject at time t. This model was fit with both the JM package in R and the %JM Macro in SAS 9.4 with results

given in Table 2.

Table 2. **CPCRA**: Current-value shared parameter joint model estimates (standard errors) with random effects standard deviations

	JM package	**%JM Macro**
Fixed Effects		
Intercept (α_0)	2.56 (0.037)*	2.52 (0.043)*
Time (α_1)	-0.043 (0.0046)*	-0.043 (0.0045)*
Time*Tx (α_2)	0.0048 (0.0066)	0.0052 (0.0064)
Random Effects		
Intercept (b_0)	0.86	0.88
Time (b_1)	0.039	0.039
Residual	0.38	0.37
Event Process (log h)		
Intercept (γ_0)	-1.41 (0.22)*	-1.40 (0.22)*
Tx (γ_1)	0.34 (0.15)	0.34 (0.15)
Association (η)	-1.13 (0.12)*	-1.13 (0.12)*
- Log Lik.	-2111	-2106

*Significant at $p < 0.05$ level.

The event process estimates were in the form of the log(hazard). As in the separate models, the estimate of treatment effect on $\sqrt{CD4}$ was in the direction of benefit of ddI over ddC, although not significantly beneficial. Additionally, treatment effect on survival was not significant. However, the coefficient was larger in magnitude compared to the coefficient from the separate survival model (0.34 vs 0.20), indicating greater harm on risk of death with ddI after accounting for the effects of $\sqrt{CD4}$ on survival. For the current-value shared parameter joint model, the ddI treatment group had on average an increase in the log(hazard) of 0.34 as compared to the ddC group after adjustment for $\sqrt{CD4}$ count. The association parameter η was significantly different from zero, indicating strong evidence of an association between the two submodels. As the estimate of η was negative, there was the indication that the true $\sqrt{CD4}$ count, $m_i(t)$, was negatively associated with the hazard of death, which was clinically sound as it has been shown that higher CD4 counts are related to better prognosis. As the true current value of $\sqrt{CD4}$ count increased by one unit, the log(hazard) decreased by 1.13 on average. Translating

these results to the bivariate outcome space, a confidence rectangle can be constructed with the 95% confidence intervals for the adjusted treatment effect of ddI on survival (γ_1) and the treatment by time effect on the $\sqrt{CD4}$ count (α_2). Using the correlation between the two treatment effects, a confidence ellipse can also be constructed. Estimates and standard errors for the longitudinal submodel and for the joint model were similar with both software methods, as seen by the overlap in Figure 5.

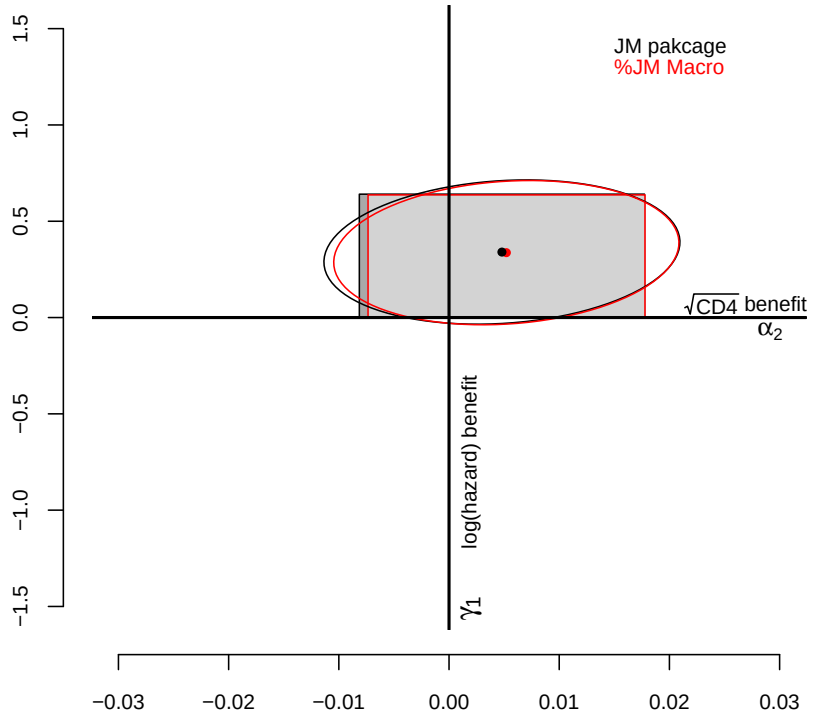


Figure 5. **CPCRA**: Current-value shared parameter joint model 95% confidence rectangles and ellipses from JM package and %JM Macro

For the CPCRA data, the slope dependent shared parameter model was given as

$$h_i(t) = \lambda_0 \exp\{\gamma_0 + \gamma_1 tx_i + \eta\, m_i'(t)\}$$

where $m_i'(t)$ was the first derivative of the expected value of the $\sqrt{CD4}$ cell count for the

ith subject at time t. This model was also fit with both the JM package in R and the %JM Macro in SAS with results given in Table 3.

Table 3. **CPCRA**: Slope dependent shared parameter joint model estimates (standard errors) with random effects standard deviations

	JM package	**%JM Macro**
Fixed Effects		
Intercept (α_0)	2.55 (0.033)*	2.52 (0.042)*
Time (α_1)	-0.045 (0.0046)*	-0.043 (0.0049)*
Time*Tx (α_2)	0.0053 (0.0064)	0.0064 (0.0064)
Random Effects		
Intercept (b_0)	0.83	0.87
Time (b_1)	0.037	0.033
Residual	0.39	0.39
Event Process (log h)		
Intercept (γ_0)	-4.63 (0.28)*	-4.15 (0.30)*
Tx (γ_1)	0.34 (0.19)	0.28 (0.16)
Association (η)	-20.3 (4.19)*	-12.6 (4.84)*
- Log Lik.	-2156	-2163

*Significant at $p < 0.05$ level.

Again, treatment was not a significant predictor of $\sqrt{CD4}$ count or survival. For the longitudinal submodel, estimates and standard errors were similar for the two software methods. For the event process portion of the joint model, estimates and standard errors were less similar but in the same in direction as the estimates from the separate and current-values shared parameter models with an increased magnitude compared to separate modeling. However, the association parameter η for the $m_i'(t)$ and survival were somewhat different in magnitude for the two software methods. They were both significant, providing evidence of association between the submodels. With a negative η, as the rate of the true slope of $\sqrt{CD4}$, $m_i'(t)$, increased, the risk of death, [log(hazard)], decreased. The differences in the estimates and standard errors for the two methods were shown visually in Figure 6 with the 95% confidence rectangles and ellipses.

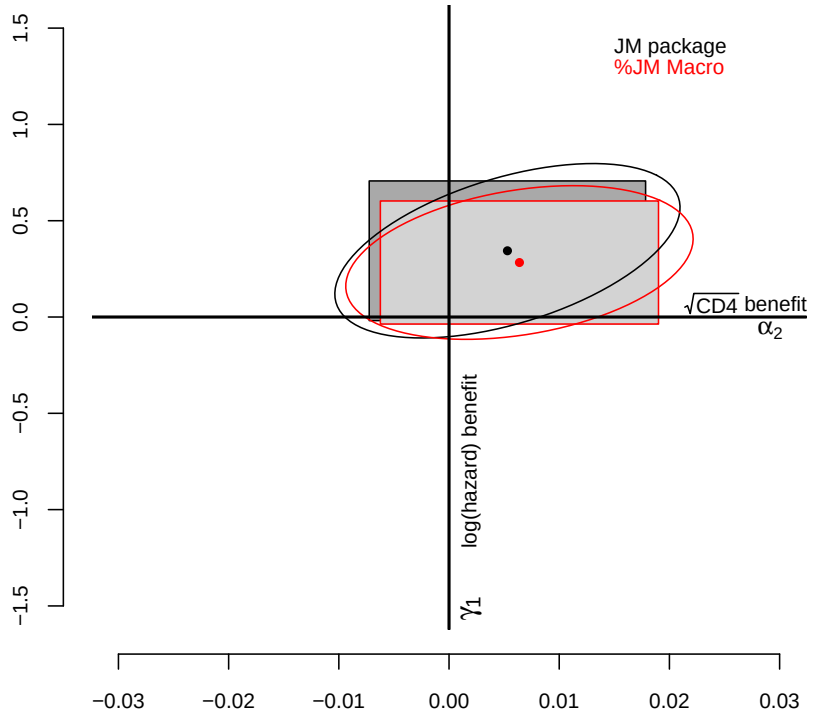


Figure 6. **CPCRA**: Slope dependent shared parameter joint model 95% confidence rectangles and ellipses from JM package and %JM Macro

The random effects coefficients shared parameter model was parameterized as

$$h_i(t) = \lambda_0 \exp\{\gamma_0 + \gamma_1 t x_i + \eta_1 b_{i0} + \eta_2 b_{i1}\}$$

where b_{i0} and b_{i1} were the random intercept and random slope parameters from the longitudinal submodel for the $\sqrt{CD4}$ count. This model was fit with both the %JM Macro and the NLMIXED procedures in SAS from Carlin's code with results given in Table 4.

Table 4. **CPCRA**: Random effects coefficients shared parameter joint model estimates (standard errors) with random effects standard deviations

	NLMIXED	**%JM Macro**
Fixed Effects		
Intercept (α_0)	2.51 (0.043)*	2.51 (0.043)*
Time (α_1)	-0.041 (0.0047)*	-0.041 (0.0046)*
Time*Tx (α_2)	0.0069 (0.0064)	0.0071 (0.0063)
Random Effects		
Intercept (b_0)	0.87	0.88
Time (b_1)	0.0362	0.037
Residual	0.37	0.37
Event Process (log h)		
Intercept (γ_0)	-3.78 (0.13)*	-3.78 (0.13)*
Tx (γ_1)	0.27 (0.15)	0.26 (0.15)
Assoc. R. Int. (η_1)	-0.95 (0.11)*	-0.95 (0.11)*
Assoc. R. Slope (η_2)	-6.96 (3.74)	-6.75 (3.60)
- Log Lik.	-2117	-2118

*Significant at $p < 0.05$ level.

Yet again, treatment was not significant in the longitudinal or the survival processes. The association parameter for the random intercept, η_1, was significant, indicating an association between the submodels, and negative, indicating the initial level of $\sqrt{CD4}$ count for an individual was negatively associated with the hazard of death. The association of the slope of $\sqrt{CD4}$ was not significant. Estimates and standard errors for the longitudinal submodel and for the joint model were almost identical for the two software methods as evidenced by the overlap of the 95% confidence rectangles and ellipses as seen in Figure 7.

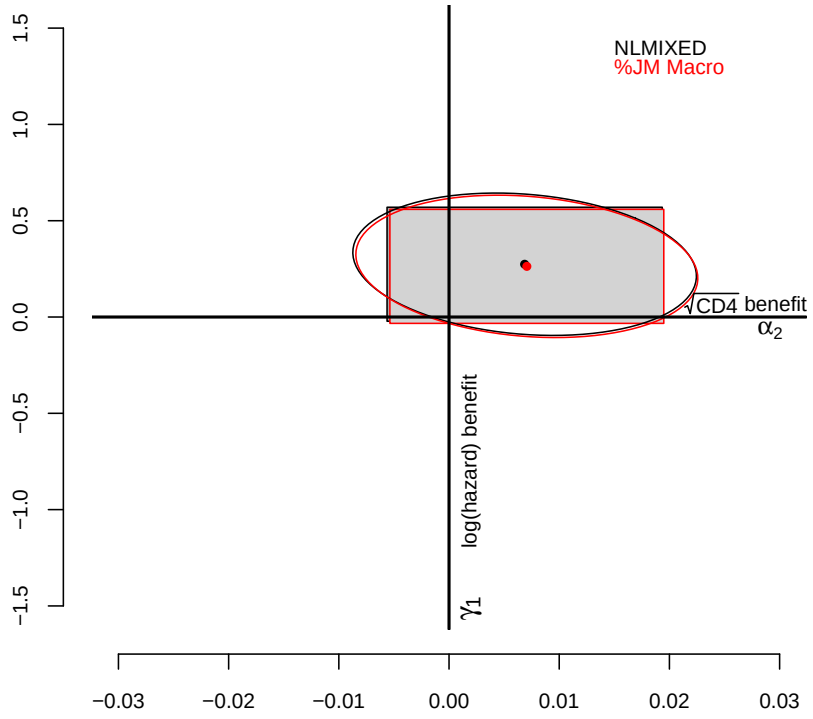


Figure 7. **CPCRA**: Random effects coefficients shared parameter joint model 95% confidence rectangles and ellipses from NLMIXED and %JM Macro

As a comparison of the ddI treatment effect parameter estimates and their standard deviations from the three different joint models and from the separate models, the 95% confidence rectangles and ellipses all from the %JM Macro in SAS for the joint models, along with the 95% confidence rectangle for the separate models, were shown in Figure 8.

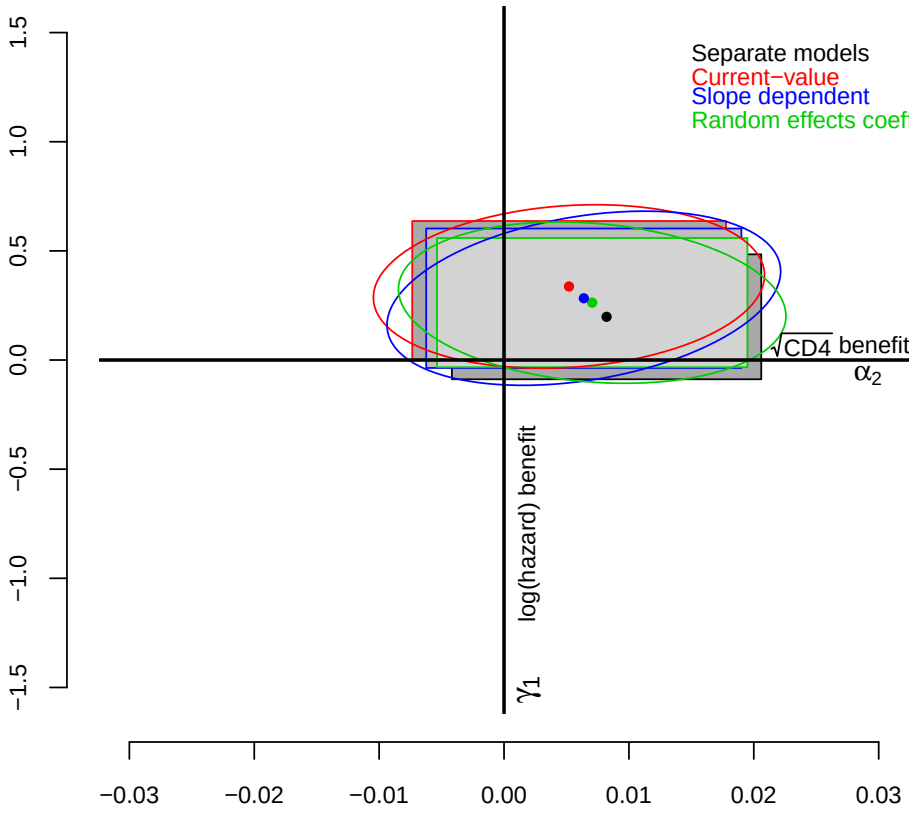


Figure 8. **CPCRA**: Comparison of shared parameter joint models with 95% confidence rectangles and ellipses from %JM Macro and 95% confidence rectangle from separate models

The estimates from the separate models and three formulations of the joint longitudinal-survival model were all in the upper right quadrant, indicating benefit of ddI on $\sqrt{CD4}$ count and harm on survival as compared to the ddC group. Additionally the 95% confidence regions all laid in this quadrant, but all regions cross both axes. Therefore all models were in agreement as to the lack of significant benefit of ddI treatment over ddC. The joint models did result in a shift of both the point estimate and 95% confidence regions upwards and slightly left, or namely in the direction of increased harm of ddI on survival and decreased benefit on $\sqrt{CD4}$.

dal-Outcomes

The dal-Outcomes trial was a F. Hoffmann-La Roche sponsored large-scale multicenter international double-blind randomized clinical trial (Schwartz et al. 2012). This dataset was chosen as an example to consider the application of joint longitudinal-survival modeling to a large clinical trial dataset yet unexplored in the joint model framework. The dal-Outcomes dataset was pertinent as stopping decisions in this trial were made based only on the effect of treatment on survival without consideration of the treatment effect on the proposed longitudinal mediator. The dal-Outcomes trial is currently relevant as it is one of three large scale cardiovascular trials examining the same class of drug that has resulted in lack of treatment benefit. Additionally, there are similar cardiovascular trials being completed for other drug classes, such as PCSK9 inhibitors, where decision criteria are based only on the supposed mediator without consideration of treatment effect on survival.

The dal-Outcomes trial was aimed at evaluating the efficacy of dalcetrapib on cardiovascular risk among patients with a recent acute coronary syndrome. Dalcetrapib is a cholesterol ester transfer protein (CETP) inhibitor that significantly raised high-density lipoprotein (HDL) levels in previous trials. Therefore, the motivation for this trial was based on a mediation mechanism where dalcetrapib would reduce the risk of cardiovascular events by raising HDL levels, as higher HDL levels have previously been shown to be associated with lower cardiovascular risk.

The study population was 15,871 patients who had an acute coronary event four to twelve weeks prior to randomization and who met inclusion criteria. Patients had measurements taken at doctor visits at randomization and during specific windows prespecified in the trial protocol. Variables such as weight, blood pressure, HDL levels, etc. were measured at doctor visits. The primary adjudicated endpoint was a survival endpoint, denoted as PCE, which was a composite of the first occurrence of death from

coronary heart disease, nonfatal myocardial infarction, unstable angina, cardiac arrest with resuscitation or ischemic stroke. The trial was stopped due to futility of dalcetrapib by the independent data and safety monitoring board at a prespecified interim analysis that included approximately 70% of the projected total number of primary endpoint events.

Again, a liner mixed effects model was used to construct the submodel for the longitudinal endpoint, HDL level (mg/dL), as

$$E[HDL_i(t)] = \alpha_0 + \alpha_1 t + \alpha_2 t * tx_i + b_{i0} + b_{i1} t \qquad G = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

where t represents time and tx_i the indicator for treatment with dalcetrapib for the ith subject. Time was measured in days with up to nine observations per subject. A random intercept and slope were also included in the model with a variance-components structure specified for the random effects variance-covariance matrix. HDL levels were normally distributed and hence not transformed. For the survival submodel for the hazard of PCE, a parametric exponential regression model was defined with constant baseline risk, giving the model

$$h_i(t) = \lambda_0 \exp\{\gamma_0 + \gamma_1 tx_i\} \quad .$$

There were 15,745 subjects that had PCE information and at least one HDL measurement giving 91,862 individual HDL observations for the cohort. There were 1,277 PCE events in this subset, for an event rate of 8.1%. Median follow up time was 31 months.

The current-value joint shared parameter model was chosen to fit the dal-Outcomes data as it was proposed that risk of cardiovascular events depends on the current HDL level. Additionally, not every subject had more than one HDL measure, and hence it would not have been possible to estimate a slope for such subjects. The joint model was

constructed as

$$h_i(t) = \lambda_0 \exp\{\gamma_0 + \gamma_1 tx_i + \eta\ m_i(t)\}$$

where $m_i(t)$ was the expected value of HDL for the ith subject at time t. This model was fit with both the JM package in R and the %JM Macro in SAS with results given in Table 5 along with results from fitting separate models.

Table 5. **dal-Outcomes**: Current-value shared parameter joint model estimates (standard errors) with random effects standard deviations

	Separate	**JM package**	**%JM Macro**
Fixed Effects			
Intercept (α_0)	47.3 (0.11)*	47.6 (0.11)*	47.4 (0.11)*
Time (α_1)	0.0030 (0.00019)*	0.0030 (0.00015)*	0.0030 (0.00017)*
Time*Tx (α_2)	0.0067 (0.00023)*	0.0068 (0.00025)*	0.0068 (0.00025)*
Random Effects			
Intercept (b_0)	13.3	12.9	13.1
Time (b_1)	0.0070	0.0071	0.0071
Residual	7.47	7.83	7.79
Event Process (log h)			
Intercept (γ_0)	-9.33 (0.040)*	-8.85 (0.11)*	-8.84 (0.11)*
Tx (γ_1)	0.045 (0.056)	0.17 (0.061)*	0.17 (0.061)*
Association (η)		-0.011 (0.0023)*	-0.011 (0.0023)*

*Significant at a $p < 0.05$ level.

With both separate modeling and joint modeling, dalcetrapib significantly increased the slope of HDL level as compared to the placebo group. Both approaches gave an estimate of similar magnitude. For both modeling approaches, one would conclude dalcetrapib had a significant effect on HDL level. However, the separate survival model and joint model would not lead to the same conclusion about the effect of dalcetrapib on the risk of PCE. With separate modeling, one would conclude that dalcetrapib had no effect on the risk of PCE as the 95% confidence interval spanned zero, even though the estimate was in the harm direction. From the joint model, one would decide that dalcetrapib significantly increased the hazard of PCE compared to placebo as the 95% confidence

interval excludes zero. The estimate of treatment effect on risk of PCE was greater in magnitude in the joint model after adjusting for the effects of HDL as compared to the separate model. From the current-value shared parameter joint model, the dalcetrapib treatment group had on average an increase in the log(hazard) of 0.17 with 95% CI (0.050, 0.29) as compared to the placebo group after adjustment for HDL level. Additionally, the association parameter η was significant and negative in the joint model, providing strong evidence that there was significant association between the submodels and that there was a negative association between HDL level and hazard of PCE. This was clinically sound as increased HDL levels have been shown to be associated with lower cardiovascular risk. For an increase of 1 mg/dL in expected current HDL level, the log(hazard) decreased by 0.011 (-0.016, -0.0065) on average. The results of Table 5 can be visualized in Figure 9 with 95% confidence regions from the joint and separate models where benefit of dalcetrapib on HDL was given by α_2 and benefit on risk of PCE by γ_1.

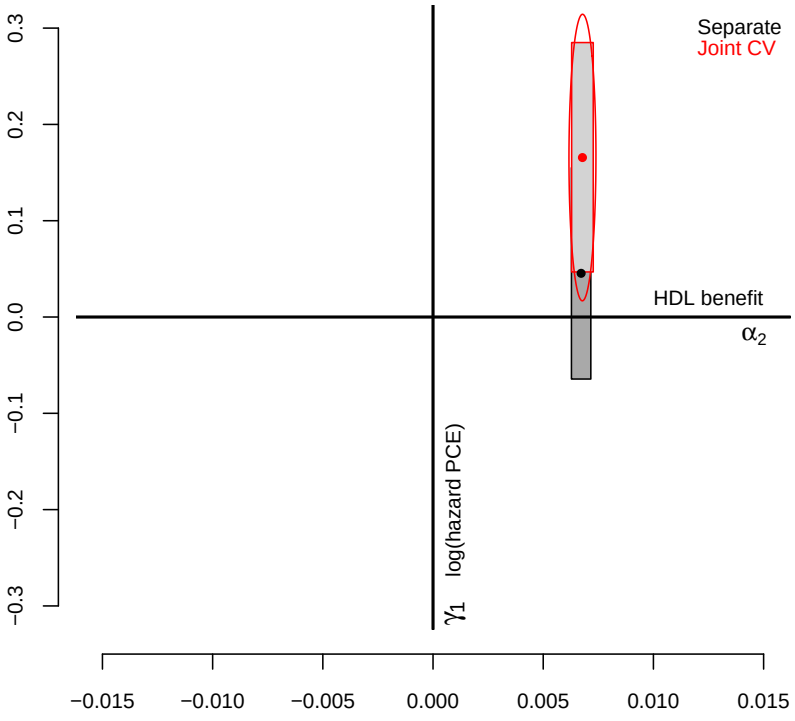


Figure 9. **dal-Outcomes**: Current-value shared parameter joint model 95% confidence rectangle and ellipse with separate models 95% confidence rectangle

CHAPTER V

SIMULATION

Methods

A simulation study was done to further compare separate modeling to joint longitudinal-survival modeling in order to examine differences in estimates and variation between the modeling methods. The classical mediation framework presented in the introduction was used as the basis for the simulation construction with the aim of simulating a basic longitudinal and survival situation. The relationship between the total effect of treatment on survival, β_1, and the direct, γ_1, and indirect effects, $\gamma_2\alpha_1$, can be described as $\beta_1 + \gamma_2\alpha_1 = \gamma_1 = b$. Here b represents the total effect of treatment on survival when α_1 equals zero, hence no treatment effect on the mediator Z. The proportion of treatment, X, effect on survival, Y, mediated through Z can be described by

$$\text{proportion mediated:} \quad \frac{\gamma_2\ \alpha_1}{\beta_1} = \frac{b - \beta_1}{\beta_1} = 1 - \frac{b}{\beta_1} \tag{V.1}$$

with signs in the desired direction for beneficial treatment effect. This was used to simulate three situations of varying proportions mediated: 25, 50 and 75% mediated. For each situation, 100 new datasets with 100 subjects per treatment group were randomly generated.

Simulations were constructed by fixing b and then determining β_1, and hence also α_1, based on the proportion mediated. The intercepts α_0 and β_0 were also fixed, as was γ_2. Six longitudinal measurements, including baseline, were simulated for the ith individual at time $t_j = 0, .., 5$ where $x_i = 1$ indicated treatment group from

$$w_{ij} = \alpha_0 + \alpha_1(x_i * t_j) + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim \text{iid } N(0, \sigma^2) \ ,$$

which were then used to estimate the slope for the ith individual from

$$z_i = \frac{\sum_{j=0}^{5}(w_{ij} - \bar{w_i})(t_j - \bar{t})}{\sum_{j=0}^{5}(t_j - \bar{t})^2}$$

where $\bar{w_i}$ was the mean of w_{ij} values for the ith individual and $\bar{t}$ was the mean of observation times. The survival outcome Y was then simulated for each individual from the mediation equation

$$y_i = \gamma_0 + \gamma_1 x_i + \gamma_2 \ z_i$$

where $\gamma_1 = b$, and due to the formulation above, which gave an approximate z for the placebo group of zero, γ_0 was set at β_0.

The Y values were then converted to survival times with the log link for exponential survival distribution with

$$y_i = log(\lambda_i)$$
$$T_i \sim \text{Exp}(1/\lambda_i)$$

where T_i were the exponential survival times. No censoring was included for the survival endpoint in order to simulate a simple case. The longitudinal observations w_{ij} for an individual were then restricted to only those which had occurred before the event time. From these restricted observations, $\hat{z}_i$ was estimated for each individual. Estimates of the treatment effect on z_i, α_1, and on y_i, β_1, were determined from separate models (denoted as $\hat{\alpha}_1^{(S)}$ and $\hat{\beta}_1^{(S)}$) along with standard errors from

$$\text{E}[z_i] = \alpha_0 + \alpha_1 x_i \tag{V.2}$$
$$\text{E}[y_i] = \beta_0 + \beta_1 x_i \tag{V.3}$$

with α_1 and β_1 determined according to the proportion mediated by equation V.1. Equation V.2 used weighted least squares regression where an individuals weight was proportional to the number of w_{ij} observations.

Treatment effect estimates and standard errors were also obtained from the joint modeling framework by fitting the restricted longitudinal observations w_{ij} and survival times T_i in a slope dependent shared parameter joint model with an exponential baseline risk as the simulation was constructed with a relationship between an individuals longitudinal slope, z_i, and event time. The estimate of treatment effect on the longitudinal observations and its standard error were taken from the fixed effects portion of the longitudinal submodel fit by the slope dependent shared parameter joint model, denoted as $\hat{\alpha}_1^{(SD)}$. The treatment effect estimate on survival time and its standard error were taken from the event process output of the slope dependent shared parameter joint model, denoted as $\hat{\beta}_1^{(SD)}$. The current-value shared parameter joint model was also fit using the simulated data giving the estimates and standard errors $\hat{\alpha}_1^{(CV)}$ and $\hat{\beta}_1^{(CV)}$. All estimates and standard errors from the separate and two joint models were averaged across the 100 datasets generated in each of the three simulations of varying proportion mediated. While reporting bivariate coverage probability would be a more meaningful summary of the simulations than averaged standard errors, 100 datasets was not sufficient to confidently estimate the coverage probability and substantially increasing the number of datasets would have been very computationally and time intensive.

Results

All simulations were carried out in R using the JM package to fit joint models. For a simulation that represented true treatment effects that were significant and in the beneficial direction as desired, the following parameter values were used: $b = \gamma_1 = -0.3$, $\beta_0 = \gamma_0 = 1$, $\alpha_0 = 5$ and $\gamma_2 = -1$. The error, ε_{ij}, was assigned to have a small σ^2. For proportion mediated of 25%, $\beta_1 = -0.4$ and $\alpha_1 = 0.1$. For 50% mediated, β_1

and α_1 equalled -0.6 and 0.3 respectively; for 75% they equalled -1.2 and 0.9 respectively. A positive α_1 indicated benefit of treatment on the mediator while the negative sign of γ_2 gave an increased benefit on survival as α_1 increased. The survival times T_i were generated such that every individual had at least two longitudinal measurements ($t = 0$ and $t = 1$) so that a slope could be estimated for each subject. Survival times were also generated such that the treatment group had a larger mean survival time than the placebo group, indicating benefit of treatment.

The average of the estimates of α_1 and β_1 and their standard errors over the 100 datasets for each scenario of proportion mediated were given in Table 6 along with the true values used in the simulation.

Table 6. **Simulation**: Estimates (standard errors*) from separate, current-value and slope dependent shared parameter joint models at varying proportions of mediation averaged over 100 datasets

Prop. Med.	**25%**	**50%**	**75%**
α_1	0.10	0.30	0.90
β_1	-0.40	-0.60	-1.20
$\hat{\alpha}_1^{(S)}$	0.079 (0.070)	0.29 (0.071)	0.87 (0.066)
$\hat{\beta}_1^{(S)}$	-0.39 (0.032)	-0.61 (0.034)	-1.21 (0.034)
$\hat{\alpha}_1^{(CV)}$	0.11 (0.021)	0.31 (0.022)	0.91 (0.020)
$\hat{\beta}_1^{(CV)}$	-0.41(0.053)	-0.58 (0.053)	-1.20 (0.046)
$\hat{\alpha}_1^{(SD)}$	0.11 (0.020)	0.30 (0.022)	0.91 (0.020)
$\hat{\beta}_1^{(SD)}$	-0.41 (0.056)	-0.59 (0.053)	-1.20 (0.050)

* Standard error equals the average of model based standard error over 100 replicates

Separate modeling, the slope dependent joint model and current-value shared parameter joint models all gave estimates of α_1 and β_1 close to truth, with little bias. Although the average of standard errors across datasets was a crude estimate of the variation in the simulations, the standard error of $\hat{\alpha}_1^{(S)}$ was consistently larger than the standard errors from $\hat{\alpha}_1^{(SD)}$ and $\hat{\alpha}_1^{(CV)}$. Additionally, the standard errors of $\hat{\beta}_1^{(SD)}$ and $\hat{\beta}_1^{(CV)}$ were consistently larger than the standard error of $\hat{\beta}_1^{(S)}$. This was not as expected; joint models were thought to have standard errors the same or smaller than separate models

for treatment effect, but this was not the case for the simulation of two example datasets. Possible reasons for this observation on the standard error of $\hat{\beta}_1$ were that both the survival and longitudinal effects were post-randomization and the computational methods used for separate and joint modeling The results showed only slight changes in the average standard errors and estimates across the varying proportions mediated. The two joint models were quite similar in their estimates, as seen by the placement of the point estimates and confidence rectangles in Figure 10.

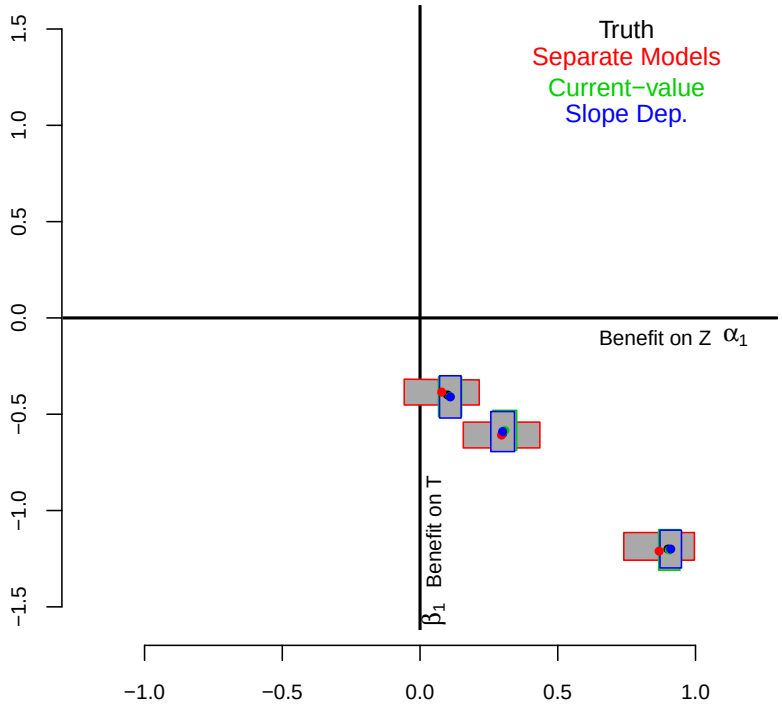


Figure 10. **Simulation**: Estimates and confidence rectangles from separate, current-value and slope dependent shared parameter joint models at varying proportions of mediation averaged over 100 datasets

The confidence rectangles shown were another somewhat crude way to display variation in the simulations. As Figure 10 showed, there was considerable overlap of the joint models' estimates and confidence rectangles after averaging over the 100 datasets for

each level of proportion mediated, while the separate model had large width confidence rectangles in the α_1 direction. A single example case from the simulations was shown in Figure 11.

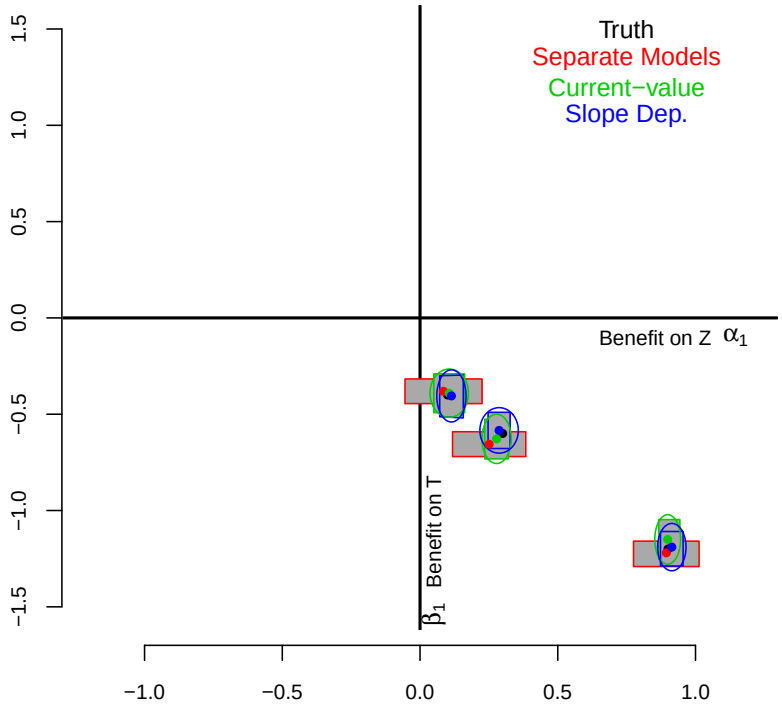


Figure 11. **Simulation**: Single case estimates and 95% confidence rectangles from separate, current-value and slope dependent shared parameter joint models at varying proportions of mediation

Displaying a single case allowed not only the plotting of 95% confidence rectangles, but also the plotting of 95% confidence ellipses from the joint models using the estimates of correlation between $\hat{\alpha_1}$ and $\hat{\beta}_1$. Again for this single case there was little bias in the estimates, although more so than the average over the 100 datasets. There was less overlap of confidence rectangles and ellipses, which was true in most single cases of the simulation. For this single case, the point estimates from the slope dependent shared parameter model were closer to truth than from the separate or current-value joint model,

which would be expected as the slope dependent model was the model under which the simulation was constructed.

CHAPTER VI

DISCUSSION

This thesis focused on constructing joint longitudinal-survival models for situations where there is a survival endpoint of primary interest as well as a companion longitudinal endpoint which may be a supposed mediator or surrogate. Three different joint models from the random effects shared parameter models framework were explored, the current-value, slope dependent and random effects coefficients shared parameter models. Three different software methods, the JM package for R, %JM Macro and the NLMIXED procedures in SAS, were used to fit example datasets.

For the CPCRA example dataset, the outcome of survival was thought to be predicted by treatment group and the surrogate longitudinal measure $\sqrt{CD4}$ count. Neither the treatment effect on survival nor on $\sqrt{CD4}$ count were significant in either the separate models or any of the joint models. Results showed an increase in magnitude of ddI treatment effect on survival with all three joint model parameterizations as compared to separate modeling, indicating an estimate of increased harm on survival of ddI after adjustment for the effects of $\sqrt{CD4}$ on survival. The association parameter for the submodels was negative and significant, providing strong evidence of an association between the submodels and a negative relationship between $\sqrt{CD4}$ count and survival. These results agreed with those previously reported on the same dataset (Guo and Carlin, 2004; Rizopoulos 2010).

The JM package and %JM Macro were independently coded, and hence, they can be used as a validation method or robustness check for one another. The two software methods were constructed based on the same framework for shared parameter joint models. Therefore, one would expect identical results from the two methods; however this was not true as seen in Tables 2 and 3. The results may be equivalent but not equal. Differences may be due to minor differences in R and SAS algorithms, convergence criteria, in optimization methods, and quadrature points for the Gauss-Hermite rule and the type

of numerical derivative used to calculate the Hessian matrix based on the score vector.

For the dal-Outcomes example dataset, the survival outcome of time to first occurrence of the PCE was supposed to be predicted by treatment group and the mediator, HDL level. Originally published results only examined treatment effect of dalcetrapib on PCE using a Cox proportional-hazards model, reporting a non-significant and futile treatment effect (Schwartz et al., 2012). However, fitting a current-value shared parameter joint model for time to PCE and HDL level resulted in a significantly harmful effect of dalcetrapib on the risk of PCE after adjustment for HDL level. There was also a significant association parameter for the submodels, indicating the appropriate use of a joint model to account for the association between the survival and longitudinal processes.

From the simulation constructed under parametric exponential survival, the estimate of the standard error for the treatment effect on the longitudinal endpoint from separate modeling was larger than from joint modeling. However, the standard error estimate for the treatment effect on survival from the joint models was larger than from the separate model, an unexpected result. Hence, there is some concern in making decision criteria in the bivariate outcome space based on joint models and these simulations can only be generalized so far. More could be done with simulations to explore this behavior and the behavior of joint modeling. A more complete evaluation would involve incorporating censoring on the survival endpoint and departures from the specified model.

One limitation in applying these joint model methods is the correct specification of the joint model as well as the longitudinal and survival submodels. Here an exponential survival process was used, although this may not apply to many cases. Additionally, the framework behind part of the motivation, including the simulation, was constructed with linear relationships between variables. This may not be true in many situations and hence results may be affected. Additionally, sources of bias and differences in standard errors were not addressed.

Further work is needed to better understand the benefits of using joint models and how

to choose the appropriate joint model for a dataset. Other shared parameter joint model constructions are available in the JM package and %JM Macro. Additionally, there are other software methods readily available for joint modeling such as the joineR package, JMBayes package and WIN-Bugs programs, with the latter two methods constructed in the Bayesian framework. Other areas of further interest are examining how joint models predict survival and applying decision making in the bivariate outcome space based on joint models to interim decisions in group sequential clinical trials.

There are some difficulties when fitting joint models. There is not a method (or methods) yet in place that are readily accepted as standard statistical approaches for fitting joint longitudinal-survival models. Additionally, analysis of large datasets with joint models can be computationally intensive as well as time demanding.

REFERENCES

Baron, A. Time-varying (or time-dependent) covariates in the cox ph model. BIOS 6646 class notes, 2014.

Chen, L. M., Ibrahim, J. G., and Chu, H. Sample size and power determination in joint modeling of longitudinal and survival data. *Statistics in Medicine*, 30(18):2295–2309, 2011.

Conlon, A. A. and Taylor, J. M. G. Surrogacy assesment using principal stratification when surrogate and outcome measures are multivariate normal. *Biostatistics*, 15(2): 266–283, 2014.

Elliot, M. R., Conlon, A. A., Li, Y., Kaciroti, N. K., and Taylor, J. M. G. Surrogacy marker paradox measures in meta-analytic settings. *Biostatistics*, 16(2):400–412, 2015.

Finkelstein, D. M. and Schoenfeld, D. A. Combining mortality and longitudinal measures in clinical trials. *Statistics in Medicine*, 18(11):1341–1354, 1999.

Fisher, L. D. and Lin, D. Y. Time-dependent covariates in the cox proportional-hazards regression model. *Annual Review of Public Health*, 20:145–157, 1999.

Gabriel, E. E., Sachs, M. C., and Gilbert, P. B. Comparing and combining biomarkers as principle surrogates for time-to-event clinical endpoints. *Statistics in Medicine*, 34 (3):381–395, 2015.

Garcia-Hernandez, A. and Rizopoulos, D. JM Macro Reference Manual, August 2015. URL `http://www.jm-macro.com`.

Gilbert, P. B. and Hudgens, M. G. Evaluating candidate principal surrogate endpoints. *Biometrics*, 64(4):1146–1154, 2008.

Guo, X. and Carlin, B. P. Separate and joint modeling of longitudinal and event time data using standard computer packages. *The American Statistician*, 58(1):1–9, 2004.

Hatfield, L. A., Hodges, J. S., and Carline, B. P. Combining longitudinal and survival information in bayesian joint models: When are treatment estimates imporved? *Biostatistics*, 0(0):1–31, 2012.

Henderson, R., Diggle, P., and Dobson, A. Joint modeling of longitudinal measurements and event time data. *Biostatistics*, 1(4):465–480, 2000.

Hoff, P. D. *A first course in Bayesian Statistical Methods*. Springer texts in statistics. Springer, 1st edition, 2009.

Hogan, J. W. and Laird, N. M. Increasing efficiency from censored survival data by using random effects to model longitudinal covariates. *Statistical Methods in Medical Research*, 7(1):28–48, 1998.

Lin, D. Y., Fleming, T. R., and De Gruttola, V. Estimating th eproportion of treatment effect explained by a surrogate marker. *Statistics in Medicine*, 16(13):1515–1527, 1997.

Parast, L., Cheng, S., and Cai, T. Incorporating short-term outcome information to predict long-term survival with discrete markers. *Biometrical Journal*, 53(2):294–307, 2011.

Parast, L., Tian, L., and Cai, T. Landmark estimation of survival and treatment effect in a randomized clinical trial. *Journal of the American Statistical Association*, 109(505): 384–394, 2014.

Philipson, P., Sousa, I., and Diggle, P. joineR - Joint modelling of repeated measurements and time-to-event data, October 2012. URL `https://cran.r-project.org/web/packages/joineR/vignettes/joineR.pdf`.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computig, Vienna, Austria, 2015.

Rizopoulos, D. JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software*, 35(9):1–33, 2010.

Rizopoulos, D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*, 67(3):819–829, 2011.

Rizopoulos, D. Package JM: Joint modelling of longitudinal and survival data, 2015. URL `{https://cran.r-project.org/web/packages/JM/JM.pdf}`.

Rizopoulos, D., Verbeke, G., and Molenberghs, G. Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival outcomes. *Biometrics*, 66(1):20–29, 2010.

Base SAS 9.4. SAS Institute Inc., Cary, NC, 2014.

Schabenberger, O. SAS NLMIXED code for joint modeling of longitudinal and survival data, March 2003. URL `http://www.biostat.umn.edu/~brad/software/LongEventExp.sas`.

Schoop, R., Schumacher, M., and Graf, E. Measures of prediction error for survival data with longitudinal covariates. *Biometrical Journal*, 53(2):275–293, 2011.

Schwartz, G. G., Olsson, A. G., Abt, M., and et al. Effects of dalcetrapib in patients with a recent acute coronary syndrome. *The New England Journal of Medicine*, 367 (22):2089–2099, 2012.

Tsiatis, A. A. and Davidian, M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*, 14:809–834, 2004.

Unknown. The em algorithm, 2006. URL `http://www.stat.ucla.edu/~yuille/courses/stat153/emtutorial.pdf`.

VanderWeele, T. J. Surrogate measures and consistent surrogates. *Biometrics*, 69(3): 561–565, 2013.

VanderWeele, T. J. and Vansteelandt, S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface*, 2:457–468, 2009.

Wang, Y. and Taylor, J. M. G. Jointly modeling longitudinal and event time data with application to acqured immunodeficiency syndrome. *Journal of the American Statistical Association*, 96(455):895–905, 2001.

APPENDIX A

CODE FOR ANALYSIS OF CPCRA DATA WITH JM PACKAGE IN R

```
#################################################################
####### FITTING JOINT MODELS IN R ####
#******** CPCRA Dataset - from JM package ************#
#define dataset
aidsjm<-JM::aids

####Separate longitudinal and survival models
#linear mixed effects model for longitudial CD4 count
fitlmeaid<-lme(sqrt(CD4)~obstime+obstime:drug,random=~obstime|patient,
data=aidsjm)
#parametric exponential survival regression for survival/death
expsep<-survreg(Surv(Time,death)~drug,data=aids.id,dist="exponential")
summary(expsep)

####Joint models with JM package##

#Does not allow longitudinal measures after event time
#lme & coxph models must be on same subjects; subjects must be
#ordered same in datasets

#define linear mixed effects model for longitudinal submodel
#can only have vc (default) random effects covar. specification
fitlmeaid<-lme(sqrt(CD4)~obstime+obstime:drug,random=~obstime|patient,
data=aidsjm)
#define cox-ph model for survival submodel
#must include x=T in model
fitsurvaid<-coxph(Surv(Time,death)~drug,data=aids.id,x=T)

#timeVar is the longitudinal time variable
#scaleWB=1 indicates exponential

#***Exponential current-value shared parameter joint model
fitjmaidexp<-jointModel(fitlmeaid,fitsurvaid,timeVar="obstime",scaleWB=1,
method="weibull-PH-GH")

#variance-covariance matrix for joint model to find correlation between
#tx effect on suvival and longitudinal endpoints
jmaidexpcov<-vcov(fitjmaidexp)
correxp<-jmaidexpcov[3,6]/(sqrt(jmaidexpcov[3,3])*sqrt(jmaidexpcov[6,6]))
coef<-fitjmaidexp$coefficients
```

```
#***Exponential slope dependent shared parameter joint model
dForm<-list(fixed=~1+drug,indFixed=c(2,3),random=~1,indRandom=2)
fitaidslope<-jointModel(fitlmeaid,fitsurvaid,timeVar="obstime",scaleWB=1,
          method="weibull-PH-GH",parameterization="slope",derivForm=dForm)
summary(fitaidslope)
fitaidslopecov<-vcov(fitaidslope)
correxpslope<-fitaidslopecov[3,6]/(sqrt(fitaidslopecov[3,3])*
sqrt(fitaidslopecov[6,6]))
```

APPENDIX B

CODE FOR ANALYSIS OF CPCRA DATA WITH %JM MACRO IN SAS

```
*Change data location accordingly;
libname my data 'C:/users/gunzbure/JMmacro';

*Change %JM macro location accordingly;
options mautolocdisplay mautosource
sasautos=('C:/users/gunzbure/JMmacro' sasautos);
options mprint nodate center;

*Change for the desired output
%let folder= C:/users/gunzbure/JMmacro/Output;

*Include necessary macros to run %JM macro
%include 'C:/users/gunzbure/JMmacro/jm.sas';
%include 'C:/users/gunzbure/JMmacro/spline.sas';
%include 'C:/users/gunzbure/JMmacro/ncspline.sas';
%include 'C:/users/gunzbure/JMmacro/calculateknotspartition.sas';
%include 'C:/users/gunzbure/JMmacro/kronrodrule15p.sas';

/*Exponential current_value shared parameter joint model*/
%JM(Data=mydata.Aids,
SubjectVar=patient,
LongiType=normal,
LongiTimeModel= linear,
LongiVar=sqcd4,
LongiTimeInter=ddi,
LongiGMatrix=vc,
EventTimeVar=time,
EventVar=death,
EventVal=1,
EventModel=exponential,
EventCovariates=ddi,
NLMIXEDOptions= ECOV,
SharedParam=current_value,
AdditionalOptions=calculateexectime);

/*Exponential slope_dependent shared parameter joint model*/
%JM(Data=mydata.Aids,
SubjectVar=patient,
LongiType=normal,
LongiTimeModel= linear,
LongiVar=sqcd4,
```

```
LongiTimeInter=ddi,
LongiGMatrix=vc,
EventTimeVar=time,
EventVar=death,
EventVal=1,
EventModel=exponential,
EventCovariates=ddi,
NLMIXEDOptions= ECOV,
SharedParam=slope,
AdditionalOptions=calculateexectime);

/*Exponential random effects coefficients shared parameter joint model*/
%JM(Data=mydata.Aids,
SubjectVar=patient,
LongiType=normal,
LongiTimeModel= linear,
LongiVar=sqcd4,
LongiTimeInter=ddi,
LongiGMatrix=vc,
EventTimeVar=time,
EventVar=death,
EventVal=1,
EventModel=exponential,
EventCovariates=ddi,
NLMIXEDOptions= ECOV,
SharedParam= bi0 bi1,
AdditionalOptions=calculateexectime);
```

APPENDIX C

CODE FOR ANALYSIS OF CPCRA DATA WITH NLMIXED IN SAS

```
proc sort data=alldata;
by patient obstime;
data all data; set all data; by patient;
last=last.patient;
run;

/*Random effects coefficients shared parameter joint model*/
proc nlmixed data=alldata;

/*all parameters not assigned starting values*/
/*explicitly assigned default=1*/
parameters r1=0.13;

/*compute log likelihood contribution of the survival data part*/
/*when the last observation of a subject is reached */
if (last) then do;
linpsurv= bs0+bs1*randgrp1+r1*u0+r2*u1;
alpha=exp(linpsurv);
G_t= exp(-alpha*t2death);
g=alpha*G_t;
llsurv=(death=1)*log(g)+(seath=0)*log(G_t);
end;  else llsurv=0;

/*Random effects variance matrix specification*/
/*ensures non-negative definite*/
v11=a11*a11;
v12=a11*a12;
v22=a12*a12+a22*a22;

/*Compute contribution of longitudinal part*/
/*Every observation in dataset makes contrib.*/
/*Conditional on random effects have independent*/
/*Gaussian contributions*/
linplong=(bl0+u0) + (bl1+u1)*obstime + bl2*obstime*randgrp1;

resid=(sqcd4-linplong);
if (abs(reside) > 1.3E100) or (s2 <1e-12) then do;
lllong=-1e20;
end; else do;
lllong=-0.5*(1.837876+resid**2/s2 +log(s2));
end;
```

```
/*Any numeric variable in the data set can be used as the*/
/*response in the MODEL statement- no effect on results*/
model last ~ general(lllong+llsurv);
random u0 u1 ~ normal([0,0],[v11,v12,v22]) subject=patient;

/*Compute median of pt specific surv. distributions*/
predict (1/alpha)*log(2) out=median;

/*Compute the var-covars. of random effects*/
/*get std. errors with delta method*/
estimate 'Var[U0]'v11;
estimate 'Cov[U0,U1] v12;
estimate 'Var[U1]'v22;
odd output ParameterEstimates=jointEstimates2;
run;
```