The Effects of Misspecification of Submodels in Joint Modeling of Repeated Measures and Time-to-Event Outcomes

by

Jason M. Mao

BS Statistics, University of California, Los Angeles, 2016

Submitted to the Graduate Faculty of the

Department of Biostatistics

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2019

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Jason M. Mao

It was defended on

December 2, 2019

and approved by

Thesis Advisor: Stewart J. Anderson, PhD, Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Ada O. Youk, PhD, Associate Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Rachel G. Miller, PhD, Research Assistant Professor, Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh

Stewart J. Anderson, PhD

The Effects of Misspecification of Submodels in Joint Modeling of Repeated Measures and Time-to-Event Outcomes

Jason M. Mao, MS

University of Pittsburgh, 2019

Abstract

Glioblastoma (GBM) is the most common form of primary brain tumor in the US. It is highly aggressive and has a median survival rate of 12 to 14 months with treatment. It has significant effects on a patient's neurocognitive functions, so consequently, patient reported outcomes (PROs) are useful for evaluating patients' physical and mental state in a way that biomarkers cannot fully capture.

Joint models, commonly used in biomedical research, combine traditional mixed models and survival analysis models, associating the longitudinal outcome with the time-to-event outcome. These models improve inferences on both types of outcomes by accounting for their underlying relationship, where events times are associated with the longitudinal outcomes.

Using data from a net-clinical benefits (NCB) sub-study of RTOG 0825, which evaluated the effects Bevacizumab on newly diagnosed GBM patients, we fit joint models to longitudinal PRO measures of symptom severity and interference with daily life and time-to-event data of GBM progression-free survival. We use these scenarios to simulate joint models where we misspecify the underlying survival and longitudinal submodels to investigate the effect of model misspecification on the association parameter that ties together the submodels.

We found that estimates of the association parameter are relatively robust to misspecification of the underlying survival distribution but not to misspecification of the assumed

trajectory of the longitudinal submodel. Individual simulations were prone to extremely biased estimates, unstable estimates, and programming errors, so further investigation is suggested.

Public Health Significance: Limited research has been done regarding the impact of misspecifying joint models. This thesis can inform methods to improve the analysis of biomarker and time-to-event data. These models, in turn, would have a public health impact when biomarkers can be used as surrogates for intervention in major health related events and thus facilitate early intervention of those events when necessary. Here, we illustrate an example confirming a result from RTOG 0825 that Bevacizumab has a negative effect on PROs in addition to investigating the association of these PROs on GBM progression.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

Glioblastoma (GBM), a type of glioma, or primary brain tumor, is the most common form of malignant brain tumor in the US. It currently has no known cure and has only seen small improvements in prognosis in recent research[1]. After diagnosis, the relative one-year survival rate of GBM patients in the US between 2000 and 2012 was 37.8% with a five-year survival rate of 5.1%. Undiagnosed cases have a median survival rate of 3 months, but this increases up to 12 to 14 months with diagnosis and treatment[2]. The current standard of care for treatment involves surgery followed by radio-chemotherapy. This thesis uses data from RTOG 0825, a phase III randomized trial of Bevacizumab among newly diagnosed GBM patients. It uses data from a net clinical benefits (NCB) substudy from this trial which focused on patient reported outcomes (PROs), including health-related quality of life and neurocognitive outcomes.

Joint models are the analytical focus of this thesis, making use of both the longitudinal and survival components of the data collected in longitudinal studies to account for the underlying association between these two types of outcomes. This is especially useful when there is nonrandom, informative dropout in the data. This thesis is a simulation study that explores the effects of either misspecifying the functional form of the longitudinal component or the underlying distribution of the parametric survival submodel on the estimation of the association parameter in joint models. We characterize the association between symptom severity and symptom interference with progression-free survival in GBM patients as the clinical example to investigate the effect of model misspecification.

In Section 2, we provide background on GBM and RTOG 0825, the motivating clinical example for this thesis. Section 3 discusses the background on joint models, the simulation of joint

models, and the methods used for the analysis. Section 4 describes the results of the analyses. Section 5 concludes this thesis with a discussion.

2.0 Clinical Background

2.1 Glioblastoma

Glioblastoma (GBM) is the most common central nervous system tumor, representing nearly 80% of malignant primary brain tumors as well as 27% of all primary brain tumors in the US. It is a highly aggressive, incurable, and deadly form of cancer, classified as a grade IV glioma by WHO classification, with a median survival rate of only 12 to 14 months after diagnosis[2]. GBM is divided into two subtypes: primary and secondary GBM. Primary GBMs appear without prior evidence of tumors and comprise the majority of GBM cases, while secondary GBMs grow from lower grade glioma. While there are no histological differences between the two, there are molecular differences that suggest the two forms develop through different pathways, so each form could benefit from differing prognosis and treatment[2]. The current standard treatment for GBM consists of maximal safe surgical resection, followed by concurrent radiotherapy with temozolomide, and then adjuvant chemotherapy with temozolomide. Surgery has shown improvements in prognosis for newly diagnosed patients, but the relative effectiveness of radiation therapy in this subgroup is more nebulous. Temozolomide is an alkylating agent that triggers apoptosis, or cell death, whose effectiveness has been reported to be correlated with the levels of methyl guanine methyl transferase (MGMT) activity. MGMT is an important DNA repair protein that reduces the effectiveness of alkylating chemotherapy against tumor cells[3]. While GBM is one of the most common forms of brain tumors, it is a relatively rare tumor with an age-adjusted incidence rate between 0.6 to 3.7 per 100,000 persons. Persons diagnosed with GBM tend to be older, with a median age of 64, tending to be higher for primary GBM (mean age of 55) than

secondary GBM (mean age of 40). It is 1.6 times more common in males than females and most common in white people compared to other ethnic groups including blacks, Asians, and Latinos[1,2]. Previous exposure to ionizing radiation is associated with increased risk of GBM, and certain genes related to allergies and the immune system are associated with decreased risk, but there is no significant evidence that lifestyle or environmental factors are associated with risk of GBM[1].

2.2 Radiation Treatment Oncology Group Study 0825 (RTOG 0825)

In this thesis, we utilize a subset of data from RTOG 0825, a phase III randomized trial of Bevacizumab for patients with newly diagnosed Glioblastoma. RTOG 0825 was funded by the National Cancer Institute and conducted as a collaboration between the Radiation Therapy Oncology Group (RTOG), the North Central Cancer Treatment Group, and the Eastern Cooperative Oncology Group. Bevacizumab is a humanized monoclonal antibody that neutralizes vascular endothelial growth factor (VEGF), an important angiogenic factor, to inhibit tumor angiogenesis. It has been approved for use in treating patients with recurrent glioblastoma but, at the time of this study, had not been studied for its effects among patients with newly diagnosed GBM[4]. Of 978 patients enrolled into the study, 621 patients were included in the final analysis. Of these 621, 309 received a placebo and 312 received Bevacizumab in addition to radiotherapy and chemotherapy with temozolomide. The study found no significant difference in overall survival between the two treatment arms, with a hazard ratio of 1.13 (95% CI: 0.93-1.37). However, there was a significant difference in progression-free survival between the placebo (median of 7.3 months) and Bevacizumab (10.7 months) groups (hazard ratio of 0.79; 95% CI: 0.66-0.94), but it is worth noting that the treatment effect varied over time. A net clinical benefits (NCB) substudy

showed that, over time, the Bevacizumab group had greater decreases in various of neurocognitive tests, including the Controlled Oral Word Association Test ($p = 0.003$) and the Trail Making Test, Part A ($p = 0.04$). It also found that the Bevacizumab had greater decreases in various symptom and health-related quality of life (QoL) outcomes such as composite symptom score ($p = 0.02$), cognitive factors ($p = 0.01$), treatment factors ($p = 0.03$), and motor dysfunction ($p = 0.02$)[4]. This thesis uses data from the NCB substudy. For these analyses we additionally restricted the data to only data from while patients are of progression-free status.

2.3 Patient Reported Outcomes (PROs)

Although the primary endpoints of the RTOG 0825 study were overall survival and progression-free survival, Patient Reported Outcomes (PROs) were analyzed in a NCB substudy of RTOG 0825. PROs provide important clinical information from patients' perspectives, allowing us to better evaluate the costs and benefits of treatments and use information that cannot be gleaned from biomedical outcomes alone, or are subjective in nature, to better interpret clinical trial results[5]. There are three components to the PROs that are assessed as part of RTOG 0825. The European Organization for Research and Treatment of Cancer (EORTC) QLQ-C30 questionnaire and BN20 module are used to assess a patient's QoL. It primarily evaluates patients' physical function, cognitive function, and affective state. The M.D. Anderson Symptom Inventory (MDASI) assesses patients' symptom severity and symptom interference with daily life, with an additional module assessing symptoms specifically related to brain tumors (MDASI-BT). Finally, neurocognitive tests, consisting of the Hopkins Verbal Learning Test-Revised, Trail Making Test Parts A and B, and the Controlled Oral Word Association test, were administered[4]. QoL can be

affected by both the disease itself as well as the treatment, with the effects of anti-tumor treatment as possibly both positive and negative. Because of GBM's nature as a brain tumor, its effect on cognitive function is especially evident, with most patients experiencing cognitive deficits and neurocognitive decline[2]. A recent study has shown that using cognitive function PROs in conjunction with tumor prognostic variables is better than using tumor variables alone to predict survival and using additional endpoints besides just survival and PFS may be useful in brain cancer clinical trials[6]. Due to the aggressiveness of GBM, losing patients to follow up results in non-ignorable missing PRO data, which we aim to use joint models in the presence of.

3.0 Statistical Methodology

3.1 Overview of Joint Models

Longitudinal studies, which are commonplace in biomedical research, are studies that follow individuals over a set period of time, collect repeated measurements of data, and typically analyze two types of outcomes: longitudinal response data and time-to-event (survival) data. Often cited examples are that of HIV and cancer research[7,8]. With HIV studies, the time to death, disease progression to AIDS, or data censoring (from study withdrawal, study closure, etc.) is recorded for each patient. In addition to baseline characteristics, repeated longitudinal measurements of biomarkers related to the immune system, such as CD4 lymphocyte count or viral load, are taken for subjects in the study. Likewise, cancer studies will record time to death or disease progression and measurements such as prostate specific antigen levels, for prostate cancer studies. Usually, research questions are constructed such that these two types of data are analyzed separately; for example, mixed effects models are used to analyze the longitudinal outcome and survival models for the time-to-event outcome, without overlap. However, in situations where the association between the two types of outcomes is of interest, joint models are useful[7,8]. If the primary focus is on the longitudinal outcome, in situations when dropout events are nonrandom and results in loss of longitudinal data, ignoring the time-to-event data can lead to biased analyses. In this situation, joint models allow us to produce valid inferences that account for this underlying relationship. If the primary focus is on the time-to-event outcome, joint models are used when we wish to analyze the effect of the time-dependent longitudinal measurements on the time-to-event outcome. Using standard survival analysis models, it is assumed that time-dependent covariables are exogeneous

and do not depend on the longitudinal outcome. When these covariates are endogenous and are related to the longitudinal outcome, such as with biomarkers and PROs, using standard survival analysis may lead to biased estimates. When the longitudinal covariates are dependent on the time-to-event outcome, then joint models are required for valid inferences[7].

3.2 Joint Model Specification

We start by defining the survival and longitudinal data we observe. We define T_i as the true event times for each of the $i = 1, \dots, n$ subjects and C_i as the underlying potential censoring times for the *i*th subject. However, we only observe $V_i = min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$, an indicator of whether we observed the event or if the subject was censored. We will define $\{y_i(t), t \geq 0\}$ as the full longitudinal response measured over all times $t \geq 0$. Because we do not observe the entire trajectory, we define $y_{ij} = \{y_i(t_{ij}), j = 1, \dots, n_i\}$ as the measurements we observe, which we collect at each of the intermittent t_{ij} time points. We use a mixed-effects model to characterize the longitudinal data while incorporating subject specific intercepts and slopes using the following equation:

$$y_i(t_{ij}) = m_i(t_{ij}) + \epsilon_{ij} = \boldsymbol{X_i^T}(t_{ij})\boldsymbol{\beta} + \boldsymbol{Z}_i^{\mathrm{T}}(t_{ij})\boldsymbol{b}_i + \boldsymbol{u}_i^T\boldsymbol{\theta} + \epsilon_{ij} \quad (1)$$

where $m_i(t_{ij})$ is the true unobserved values of the longitudinal response at times t_{ij} and the error term $\epsilon_{ij} \sim N(0, \sigma^2)$ represents measurement error, which we assume are mutually independent and independent of the random effects $\boldsymbol{b}_i$. We define $\boldsymbol{X}_i$ as a design matrix for the fixed effect $\boldsymbol{\beta}$ and $\boldsymbol{Z}_i$ for the random effects $\boldsymbol{b}_i$, with distributed as multivariate normal $\boldsymbol{b}_i \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$[7]. Baseline covariates are represented with vector $\boldsymbol{u_i}$ with a corresponding vector of regression coefficients $\boldsymbol{\theta}$ to delineate

from the covariates we measure over time in $\boldsymbol{X_i}$, although some authors include it within the $\boldsymbol{X_i}$ term[8]. For longitudinal outcomes that show non-linear trajectories, flexible formulations for $m_i(t)$ are preferred, using functions of time $\boldsymbol{t}$ expressed as high-order polynomials or splines. Splines are considered the preferred way to model highly nonlinear trajectories because they have better numerical properties and, due to their local nature, avoid possible problems associated with the global nature of polynomials[7]. Approaches utilizing cubic b-splines have been proposed by several authors[9,10]. For example, Brown (2005) proposes a Bayesian hierarchal model that defines

$$Y_{ij} = \psi(t_{ij}) + \epsilon_{ij} = \sum_{k=1}^{q} \beta_{ik} B_k(t_{ij}) + x_i'\alpha + \epsilon_{ij} \quad (2)$$

where $\beta_{ik} \sim N(b_{0k}, V_{0k})$. The summation term is a random effect curve with a q-dimensional basis for spline functions on $[0, T]$ and the $x_i'\alpha$ term accounts for the effect of the baseline covariates. This model can also be extended to the multivariate case of longitudinal outcomes[10]. Another alternative framework uses models with the form

$$y_i(t) = m_i(t) + U_i(t) + \epsilon_i(t) \quad (3)$$

where $U_i(t)$ is a mean-zero stochastic process. Possible specifications of $U_i(t)$ include as an integrated Ornstein-Uhlenbeck process or as a stationary Gaussian process, allowing trends to vary over time and accounting for biological fluctuation about a smooth trend[8]. For example, in Henderson (2000), the following specification is proposed:

$$U_{1i}(t) = W_{1i} + W_{2i}t \quad (4)$$

$$U_{2i}(t) = \gamma_1 W_{1i} + \gamma_2 W_{2i} + \gamma_3(W_{1i} + W_{2i}t) + W_{3i}$$

with the subscript *i* left out. In this, (W_1, W_2) are latent zero-mean bivariate normal variables and $W_3 \sim N(0, \sigma_3^2)$, independent of (W_1, W_2). The parameters γ_1, γ_2, and γ_3 measure the association

between W_1 and W_2, while W_3 is a frailty term[11]. For our formulation, we will focus on $m_i(t)$ as a polynomial function of time, a simple example being $m_i(t) = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t + \beta_2 t^2$.

Finally, we define the survival submodel using the relative risk formulation of the proportional hazards model:

$$h_i(t \mid M_i(t), v_i) = \lim_{dt \to 0} Pr\{t \leq \mathrm{T_i} < t + dt \mid T_i \geq t, M_i(t), v_i\}/dt$$

$$= h_0(t) exp\{\eta^T v_i + \alpha m_i(t)\} \quad (5)$$

where $M_i(t) = \{m_i(s), 0 \leq s < t\}$ is the true, complete history of the unobserved longitudinal process up to time *t*. The baseline hazard function is denoted as $h_0(\cdot)$ and v_i denotes a vector of baseline covariates that may or may not be the same as $\boldsymbol{u}_i^T$ in (1) with corresponding coefficient vector η. The Cox Proportional Hazards model is the most popular choice of model to use for the survival submodel. In it, the baseline hazard function is left unspecified and is estimated nonparametrically while η is estimated using a partial likelihood function. This is known as the semi-parametric approach[7]. The method this thesis focuses on is the parametric approach that specifies a known parametric function, such as the exponential function, for the baseline hazard function. The parameter α is the measure of association between the longitudinal outcome to the time-to-event outcome risk. Previous studies have evaluated the potential of using the longitudinal outcome as a surrogate for the time-to-event outcome with the conditions that if the treatment had an effect on the time-to-event, the treatment had an effect on the longitudinal outcome, and v_i in (5) includes the treatment indicator, then the risk of the event based on the longitudinal outcome should be independent of treatment, so $m_i(t)$ can serve as a surrogate[8]. Alternative specifications for the survival submodel include the rate of change structure that uses $\alpha m_i'(t)$ instead of $\alpha m_i(t)$ in (5) or the accelerated failure time framework. Though we focus on continuous longitudinal

outcomes in this thesis, joint models can be extended to categorical and count data within the generalized linear mixed model framework[7].

Initial approaches to fit joint models were based on two-stage models; in the first stage random effects are estimated using growth curve models, and in the second stage the estimates are substituted into the partial likelihood of the Cox model which is then maximized. Now, the primary estimation method for joint models uses maximum likelihood estimation. The likelihood function for the parameters of interest is

$$\prod_{i=1}^{n} \int [h_0(V_i) exp\{\eta^T v_i + \alpha m_i(V_i)\}]^{\delta_i} exp\left[-\int_0^{V_i} h_0(u) exp\{\eta^T v_i + \alpha m_i(u)\} du\right]$$

$$\times \frac{1}{\{(2\pi\sigma\}^2)^{m_i/2}} \exp\left[-\sum_{j=1}^{m_i} \frac{Y_i(t_{ij}) - m_i(t_{ij})^2}{2\sigma^2}\right] p(\alpha_i|\nu_i; \delta) d\alpha_i \qquad (6)$$

where $p(\alpha_i|\nu_i; \delta)$ is assumed multivariate normal.

Maximization of the log-likelihood function can be done using the Expectation-Maximization (EM) algorithm or the Newton-Raphson algorithm, though the EM algorithm is preferred due to some of the parameters having closed-form estimates[7].

3.3 Effects of Model Misspecification

In this thesis we are interested in the effect model misspecification has on the estimation of joint model parameters. Our primary interest is in the effects of misspecifying the forms of the $m_i(t)$ time polynomial and the underlying baseline hazard function $h_0(\cdot)$ on the association parameter α in the survival submodel of the joint model, a topic that has received limited attention. Gail (1984) shows that, in the analysis of uncensored survival data, parametric proportional

hazards models give unbiased estimates of treatment effects regardless of the underlying survival distribution. However, when censoring is present, estimates are biased. When needed covariates are omitted, the estimates of treatment effects are also biased, but if the exponential model is used, less bias is observed than if using the Cox proportional hazards model. In proportional hazards models, when covariates are omitted, estimates of the regression parameters are asymptotically biased towards zero. The degree of bias is expected to be small, unless if the true value is large[12]. Asymptotic variances from the model with omitted covariates are also smaller than those from the true model[13]. The effect of misspecifying the hazard function in parametric proportional hazards models will vary depending on the true underlying distribution of the survival data. In a study modeling against a single covariate, the size and power of hypothesis tests on the parameters of misspecified parametric models is generally outperformed by Cox models, with exponential models performing particularly poorly except in the case of no misspecification[14]. Kwong (2003) states that there is merit to fitting parametric survival models, as models are generally robust to misspecification if coefficients are small or if survival times are long, including in the presence of influential observations. However, this less true when hazard rates are rapidly increasing or decreasing[15]. Misspecifying the underlying survival distribution can also lead to models with local overfitting and increased bias at the lower and upper percentiles compared to at the median[16,17]. Correctly specified and asymptotically well-fitting parametric models should give efficient parameter estimates[16], but the potential trade-off for misspecification of the underlying survival distribution leading to highly biased estimates is undesirable. This issue is exacerbated when there is a large amount of censoring present in the data[14,17]. As such, some authors caution against the use of parametric proportional hazards models due to the possibility of producing very biased estimates from misspecified models and misleading results for the shape and scale parameters of

the hazard function unless if prior knowledge suggests using a specific parametric model. They suggest using a Cox proportional hazards or accelerated failure time models as they are more robust to misspecification of the underlying survival distribution[16].

3.4 Simulation of Joint Models

Bender (2005) provides a general framework for simulating survival data from parametric proportional hazards models[18]. First, the hazard function *h(t)* of proportional hazards models is given as

$$h(t|X) = h_0(t)exp(X\beta) \tag{7}$$

where $h_0(t)$ is the baseline hazard function of a parametric distribution,

$$H(t|X) = H_0(t)exp(X\beta), \qquad where\ H_0(t) = \int_0^t h_0(u)du \tag{8}$$

describes the cumulative hazard, and the survival function, *S(t)*, and cumulative distribution function, *F(t)*, are as follows:

$$S(t|X) = exp[-H(t)] \quad and \quad F(t|X) = 1 - exp[-H(t)]. \tag{9}$$

Here, t represents time, X the baseline covariate vector, and β the associated regression coefficients. Cumulative distribution functions of variables follow a standard uniform distribution, denoted here as $U \sim Unif(0,1)$, and it follows that $1 - U$ is also distributed as $Unif(0,1)$. From this, letting T be the event time, we have

$$U = H(t|X) = exp[-H_0(T) \times exp(X\beta)] \sim Unif(0,1) \tag{10}$$

which we can then solve directly for T as long as $H_0(t)$ can be inverted. So, using the following equation

$$T = H_0^{-1}[-log(U) \times exp(-X\beta)] \quad (11)$$

we only need to generate random variables from $U \sim Unif(0,1)$ in order to easily simulate survival data. Using the simple example of the exponential distribution, the baseline hazard function is $h_0(t) = \lambda$, giving $H_0(t) = \lambda t$ and $H_0^{-1}(t) = \lambda^{-1}t$. Plugging this into (11) gives us

$$T = \lambda^{-1}[-log(U) \times exp(-X\beta)] = -\frac{log(U)}{\lambda \times exp(X\beta)} \quad (12)$$

However, when the equations become more complex, such as in the context of joint models, this method cannot be directly applied to simulate survival data. Crowther (2013) describes methods to extend Bender's framework to cover situations such as when there is a complex baseline hazard function, time-dependent effects, time-varying covariates, and random effects, such as in joint modelling[19]. In these scenarios, if $H_0(t)$ does not have a closed form solution or if *T* cannot be solved for analytically, numerical integration and iterative root finding methods are required to solve for the simulated times *T*. In short, if in equation (8), $h_0(t)$ is a complex function of *t*, or if $exp(X\beta)$ is a function of *t*, we must use numerical integration to calculate $H(t|X)$ and then use iterative root finding methods to find the simulated time *T* that solves equation (10).

4.0 Results

4.1 RTOG 0825 Data Analysis

Our analyses use a 477 patient subset from the NCB substudy of the RTOG 0825 clinical trial, only collecting data from patients who had consented to participating in the NCB component of the study and whose tumors had not yet progressed. Among these patients, 229 (48.0%) were assigned to receive the placebo and 248 (52.0%) were assigned to receive the Bevacizumab treatment. The baseline characteristics of MGMT status and RPA class, a prognostic classification for GBM patients, were well balanced between the two treatment arms. These covariates are significantly associated with GBM prognosis, but in Gilbert (2014) were not found to have a significant differential effect on treatment effect or to change the treatment effect when adjusted for, and we will not be using them in our analyses[4]. In our data subset, disease progression occurred in 191 (83.4%) of placebo group patients and 201 (81.0%) of Bevacizumab group patients. The median survival time in the placebo group was 7.3 months (95% CI: 5.7 to 8.4) compared to 10.8 months (95% CI: 10.0 to 12.4) in the Bevacizumab group, with a hazard ratio of 0.82 (95% CI: 0.67 to 1.00, $p = 0.047$ by the log-rank test). When we plot the Kaplan-Meier curve (Figure 1), we see that early on the Bevacizumab has less progression events than the placebo group. However, after approximately 1.5 years the curves cross, indicating that the proportional hazards assumption of the hazard ratio has been violated.

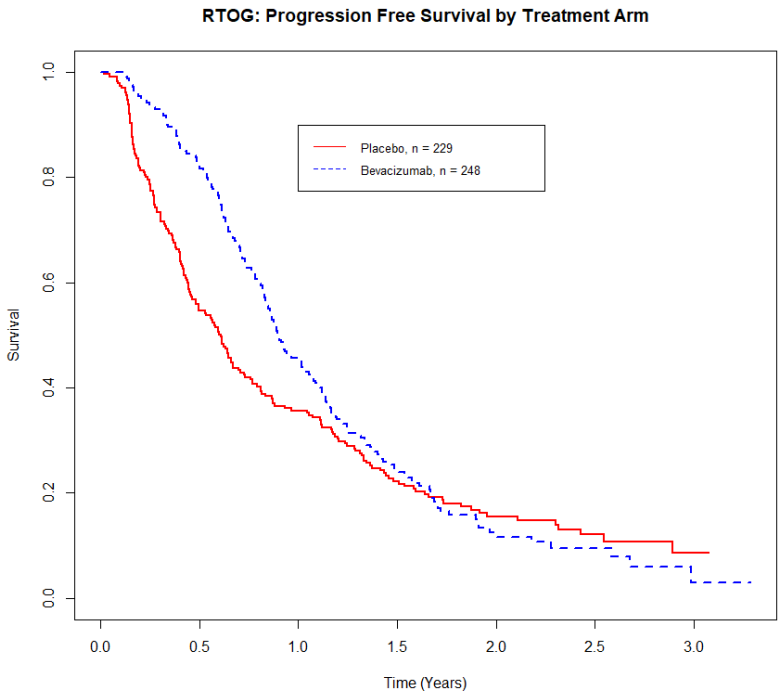

Figure 1 Progression-Free Survival Kaplan-Meier Plot

The NCB substudy in Gilbert (2014) identifies several PROs that are found to have a significant differential deterioration over time between the Bevacizumab and placebo treatment groups. PRO data is collected during assessments at 6 time points: 0, 6, 10, 22, 34, and 46 weeks after randomization. For our analyses, we will focus on two MDASI-BT outcome measures: the composite symptom score and the composite symptom interference score. These are measures of the symptom severity and the symptom interference with daily life, scored on a 0-10 scale with 0 representing no severity and no interference and 10 representing extreme severity and complete interference. In total, 13 severity items are averaged for a composite symptom score and 6 interference items are averaged for the composite symptom interference score (Appendix B). These scores are highly right-skewed, so we will use the square roots of the scores in our models to limit issues that may stem from modeling skewed data. As such, subsequent references to the symptom severity and interference scores will refer to their square roots.

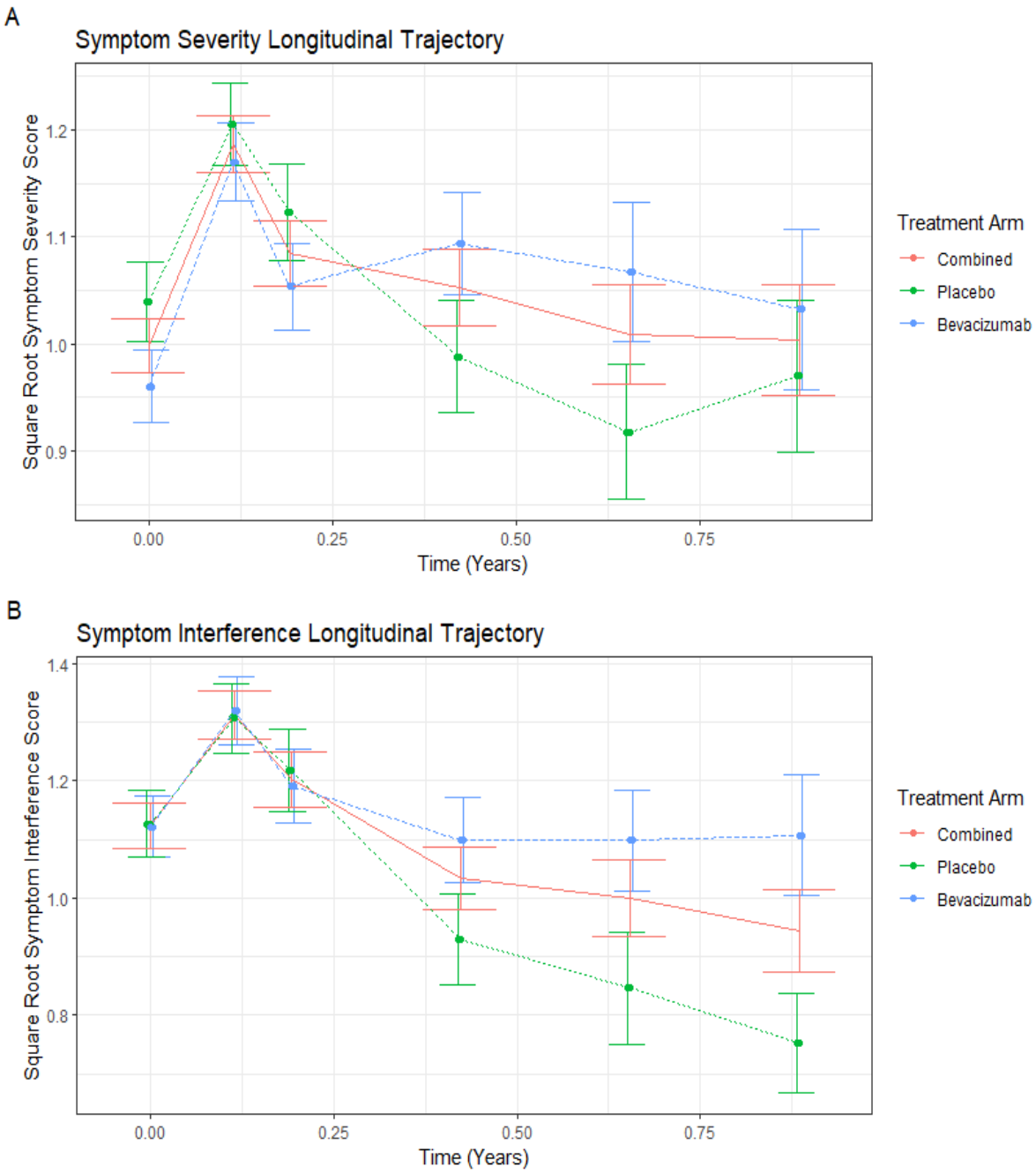

Figure 2 Observed Longitudinal Trajectories of Symptom Severity and Interference

In the plots of the symptom severity (Figure 2A) and interference (Figure 2B), the two groups start very close, but after about 0.25 years (3 months), the curves separate, with the Bevacizumab tending towards higher scores on average than the placebo group, meaning that the

Bevacizumab is experiencing worse symptom severity and symptom interference. Observing the approximately cubic shape to the symptom severity curve, we fit the following longitudinal submodel:

$$y_i(t_{ij}) = m_i(t_{ij}) + \epsilon_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 Trt_i + \beta_5 t * Trt_i + \epsilon_{ij} \quad (13)$$

where Trt_i is a treatment group indicator, time in the scale of years, and

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_0^2 & \rho\sigma_0\sigma_1 \\ \rho\sigma_0\sigma_1 & \sigma_1^2 \end{bmatrix} \right) \quad , \qquad \epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \quad (14)$$

which is a model with fixed quadratic and cubic powers of time, fixed treatment effect and treatment interaction with linear time, and random intercept and linear slope. We fit the following survival submodel with an exponential baseline hazard function

$$h_i(t) = \lambda exp\big(\eta * Trt_i + \alpha m_i(t)\big) \quad (15)$$

where $m_i(t)$ is from (13). For the symptom interference joint model, we fit a similar submodel without the cubic time term:

$$y_i(t_{ij}) = m_i(t_{ij}) + \epsilon_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t + \beta_2 t^2 + \beta_4 Trt_i + \beta_5 t * Trt_i + \epsilon_{ij} \quad (16)$$

with random effects and measurement error specified just as in (14) and the survival submodel specified as in (15), using the $m_i(t)$ from (16). Though standard survival parametric regression suggests that the survival distribution may be better fit as Weibull, these joint models fail to run with a Weibull survival submodel on our data, so we will use the exponential baseline hazard for our joint models. Standard mixed models of the symptom interference model also suggest modeling with cubic time, but joint models with the cubic term fail to converge, so we fit a quadratic model. Analyses are performed using the JM package in R (see Appendix A).

Results for the symptom severity model described above is in the following Table 1 and symptom interference in Table 2. In both the symptom severity and interference joint models the

Table 1 Symptom Severity Score Joint Model Results

Parameter	Coefficient	Std. Error	p-value
Longitudinal Submodel			
Intercept β_0	1.0564	0.0323	<0.0001
Time β_1	1.1738	0.2166	<0.0001
Time^2 β_2	-3.4508	0.6201	<0.0001
Time^3 β_3	2.3895	0.4710	<0.0001
Treatment β_4	-0.0677	0.0425	0.1115
Time*Treatment β_5	0.2376	0.0879	0.0069
Survival Submodel			
Treatment η	-0.0648	0.1013	0.5525
Association α	-0.0003	0.0001	0.0029
Log(Scale) $log(\lambda)$	-1.2869	0.0745	<0.0001
Random Effects Variance-Covariance			
σ_0	0.3826		
σ_1	0.3517		
ρ	0.2495		
Measurement Error			
σ_ϵ	0.3489		

Table 2 Symptom Interference Score Joint Model Results

Parameter	Coefficient	Std. Error	p-value
Longitudinal Submodel			
Intercept β_0	1.2451	0.0486	<0.0001
Time β_1	-0.5564	0.1289	<0.0001
Time^2 β_2	0.2556	0.1011	0.0115
Treatment β_3	-0.0144	0.0666	0.8284
Time*Treatment β_4	0.3710	0.1332	0.0053
Survival Submodel			
Treatment η	-0.0461	0.1023	0.6520
Association α	-0.0342	0.0153	0.0254
Log(Scale) $log(\lambda)$	-1.2299	0.0782	<0.0001
Random Effects Variance-Covariance			
σ_0	0.5834		
σ_1	0.4662		
ρ	-0.1328		
Measurement Error			
σ_ϵ	0.5763		

effects of time and the interaction between time and treatment were significant, but not the treatment main effect, in the longitudinal submodel. The interaction coefficient in both models indicates that the Bevacizumab group has a more quickly increasing slope in symptom scores than the placebo group. In the survival submodel, neither joint model found the treatment effect to be significantly associated with progression-free survival. Both models found a significant association between the longitudinal outcome and survival, with α = -0.0003 (p = 0.003) in the symptom severity model and α = -0.0342 (p = 0.025) in the symptom interference model. Interestingly, the models seem to indicate increased (worse) symptom scores are associated with a reduction in hazard rate. In the symptom severity joint model, including cubic time leads to rapidly increasing values of the outcome, so the association parameter is sensitive to this and is very small compared to in the symptom interference joint model's association parameter. These results will be used to inform our simulations.

4.2 Simulation Results

Our first simulation is based on the symptom severity joint model, specifying the assumed longitudinal trajectory as in (13) and the survival submodel as in (15), using the parameter values of $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \sigma_0, \sigma_1, \rho, \lambda, \eta,$ and σ_ϵ from Table 1 to simulate data and fit joint models. We then misspecify the survival submodel to have a Weibull underlying distribution, using

$$h_0(t) = \lambda^\gamma \gamma t^{\gamma-1} \tag{17}$$

as the baseline hazard function. Our second is based on the symptom interference joint models, with (16) and (15) as the assumed longitudinal and survival model specification, using the parameters from Table 2. We then misspecify the longitudinal component as

$$y_i(t_{ij}) = m_i(t_{ij}) + \epsilon_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t + \beta_4 Trt_i + \beta_5 t * Trt_i + \epsilon_{ij} \qquad (18)$$

without a quadratic time term. Longitudinal outcomes are simulated at up to 6 time points: 0, 0.2, 0.4, 0.6, 0.8, and 1.0 years, evenly spaced and covering a similar time period as the original data. Treatment group membership is simulated from a *Bernoulli(0.5)* distribution. A 5% chance of random missingness in the longitudinal data is generated from a uniform $U(0,1)$ distribution. Survival data is simulated as described in Section 3.3, with disease progression times and censoring times simulated according to the same hazard function, as we assume censoring is informative, selecting the minimum of the two as the observed event. Longitudinal measurements simulated after the observed event time are deleted as well as patients with zero longitudinal data. Administrative censoring is applied at 3 years. For each model, 500 simulations are generated with longitudinal and survival outcomes for 200 patients in each simulation (Appendix A).

Results for the simulated symptom severity joint models are presented in Table 3, where we display the bias, percent bias, coverage probability (CP), and mean square error (MSE) of the estimates of α from the simulations. For the correctly specified exponential baseline hazard model, of 500 simulations, 45 (9%) had non-positive-definite Hessian matrices and gave parameter estimates that cannot be used. Another 122 (24.4%) entirely failed to run, giving fatal errors, leaving N = 333 (66.6%) simulations to analyze. In the incorrectly specified Weibull baseline hazard model, 42 (9.4%) had non-positive-definite Hessian matrices and 153 (30.6%) encountered fatal errors, leaving N = 305 (61%) simulated models to analyze. In the longitudinal submodel, we note a moderate amount of bias in the time parameters of both the correctly and incorrectly specified models and less so in estimates of the treatment effects (Table 3). The Weibull model tends to be less biased but have lower coverage probability than the exponential model. In the survival submodel, both joint models are moderately biased in the estimate of the treatment effect

and extremely biased with respect to the association parameter α, though the correctly specified model performs slightly better. The simulated estimates of the variance-covariance matrices and σ_ϵ seem to perform well in both models. However, as seen in Figure 3A, there are extreme observations heavily influencing the mean of the simulated alpha parameters from both joint models. In Figure 3B, we see trimming the 20% of the simulations based on the alpha parameter

Table 3 Misspecified Survival Submodel Simulation Results

Parameter	**Truth**	**Bias**	**% Bias**	**CP**	**MSE**	**Bias**	**% Bias**	**CP**	**MSE**
		Exponential (N = 333)				**Weibull (N = 305)**			
Longitudinal Submodel									
Intercept β_0	1.056	0.001	0.11	91.9	0.016	-0.000	-0.04	84.4	0.017
Time β_1	1.174	-0.268	-22.84	74.5	1.944	-0.250	-21.31	68.8	2.072
Time^2 β_2	-3.451	0.773	-22.39	72.7	16.312	0.729	-21.13	67.0	17.427
Time^3 β_3	2.390	-0.531	-22.24	71.8	7.700	-0.503	-21.03	66.7	8.230
Treatment β_4	-0.068	0.003	-5.04	94.0	0.008	0.003	-5.01	85.9	0.009
Time* Treatment β_5	0.238	-0.005	-2.02	96.4	0.008	-0.005	-2.20	88.6	0.008
Survival Submodel									
Treatment η	-0.065	-0.020	30.62	93.4	0.057	-0.023	35.55	85.6	0.056
Association α	-0.0003	-0.083	31668	65.2	4.009	-0.091	34649	58.9	4.377
Random Effects Variance-Covariance									
σ_0	0.383	-0.005	-1.3		0.001	-0.005	-1.2		0.001
σ_1	0.352	-0.015	-4.3		0.005	0.017	-4.9		0.006
ρ	0.249	0.019	7.6		0.026	0.004	1.5		0.030
Measurement Error									
σ_ϵ	0.349	0.009	2.7		0.010	0.009	2.5		0.009

still leaves us with a substantial right tail. In Figure 3C, trimming by 40% gives a better picture of the distribution of the simulated alphas near 0. We see that the simulated alphas are biased positively, away from the true parameter. In Table 4, when examining the median and trimmed means of the simulated alphas, while still moderately biased, it is significantly less biased than the untrimmed mean. Here the misspecified Weibull model is less biased and generally performs better than the correctly specified exponential model.

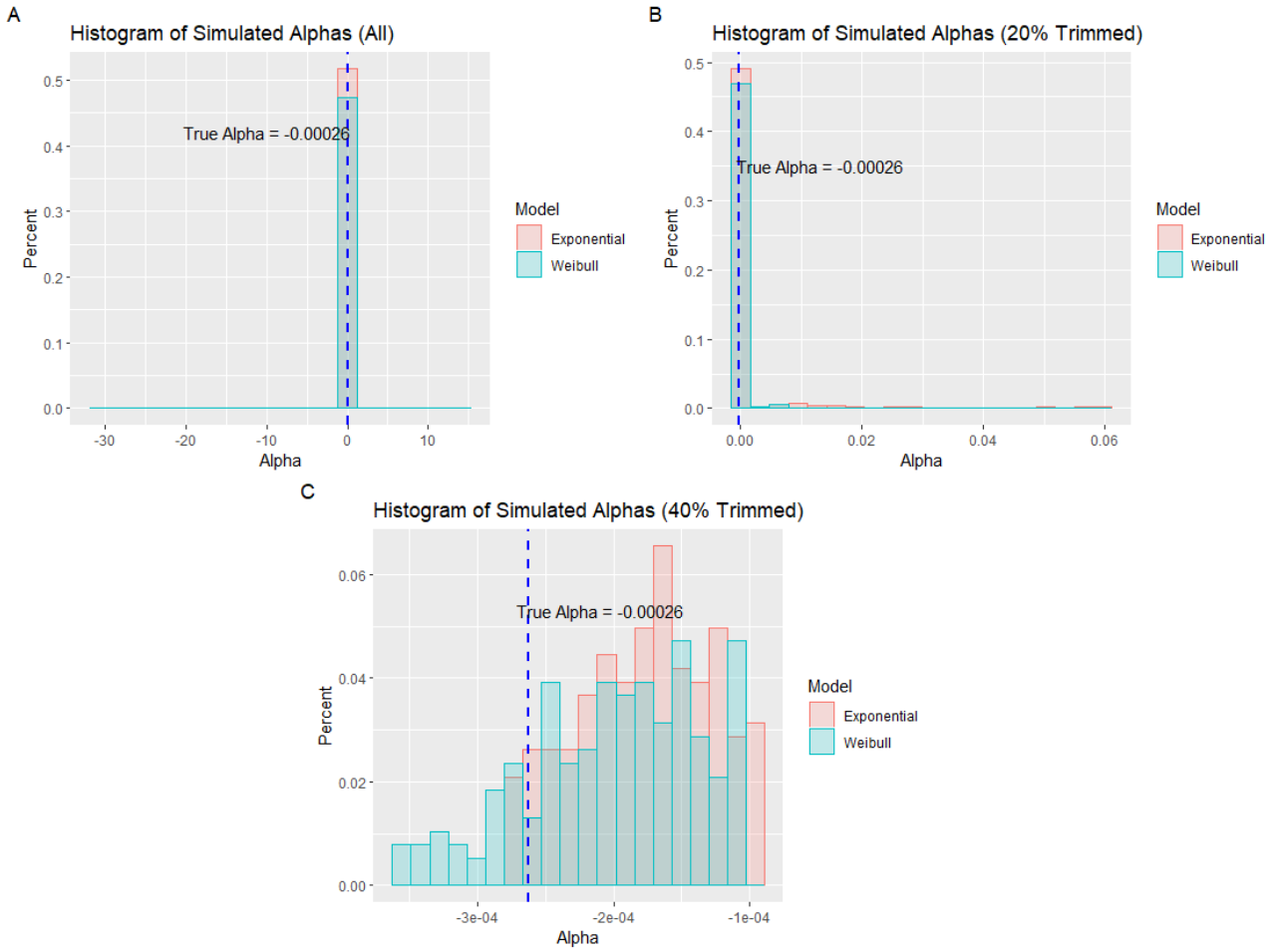

Figure 3 Histogram of Alphas from Simulation 1

Table 4 Trimmed and Median Alphas from Simulation 1

	Truth (10^{-2})	**Bias (10^{-2})**	**% Bias**	**CP**	**MSE (10^{-6})**	**Bias (10^{-2})**	**% Bias**	**CP**	**MSE (10^{-6})**
		Exponential				**Weibull**			
20% Trimmed	-0.026	0.140	-532.3	68.5	46.308	0.015	-58.0	71.2	0.737
40% Trimmed		0.008	-32.5	67.2	0.010	0.006	-23.9	76.2	0.008
Median		0.009	-34.6			0.007	-27.0		

Results for the simulated symptom interference joint models are presented in Table 5. There are significantly fewer errors in these joint models. In the correctly specified quadratic model, 3 (0.6%) had non-positive definite Hessian matrices and 13 (2.6%) encountered fatal errors,

leaving 484 (96.8%) to analyze. In the misspecified linear model, there are 4 (0.8%) with non-positive-definite Hessian matrices and 30 (6%) that ran with fatal errors, leaving 469 (93.8%) simulations for analysis. Here, we see that the moderate to significant amount of bias present in

Table 5 Misspecified Longitudinal Submodel Simulation Results

Parameter	**Truth**	**Bias**	**% Bias**	**CP**	**MSE**	**Bias**	**% Bias**	**CP**	**MSE**
		Quadratic (N = 484)				**Linear (N = 469)**			
Longitudinal Submodel									
Intercept β_0	1.245	-0.007	-0.6	92.6	0.006	-0.015	-1.2	91.7	0.008
Time β_1	-0.556	0.114	-20.5	58.7	0.082	0.237	-42.5	34.5	0.072
Time^2 β_2	0.256	-0.121	-47.4	53.3	0.076				
Treatment β_3	-0.014	-0.004	25.6	94.0	0.012	0.020	-138.3	89.8	0.014
Time* Treatment β_4	0.371	0.003	0.9	95.2	0.020	-0.065	-17.5	79.7	0.035
Survival Submodel									
Treatment η	-0.046	-0.027	57.6	95.0	0.062	-0.180	390.5	81.7	0.148
Association α	-0.034	0.042	-124.4	50.4	0.006	0.158	-463.2	44.3	0.088
Random Effects Variance-Covariance									
σ_0	0.583	-0.007	-1.2		0.002	0.003	0.6		0.027
σ_1	0.466	-0.049	-10.5		0.019	-0.149	-32.0		0.056
ρ	-0.136	0.020	-15.2		0.037	0.051	-38.0		0.350
Measurement Error									
σ_ϵ	0.576	0.002	0.400		0.000	0.008	1.4		0.001

the correctly specified simulations is exacerbated in the misspecified simulation joint models for all parameters. We also observe that coverage probability for the time parameters is fairly low in the quadratic simulations and even lower in the linear simulations, although the simulated estimates of the treatment parameters performed relatively well. Significant bias of the association parameter is also present here, though not nearly to the extent in the symptom severity simulations. As seen in Figure 4A, like the symptom severity model simulations, there are large values of the simulated alphas affecting the mean bias, though not as extreme as in the previous simulations. Figure 4B shows the 20% trimmed histogram of alpha parameter estimates. We observe that the misspecified linear model has a longer right tail, though both simulations are right skewed.

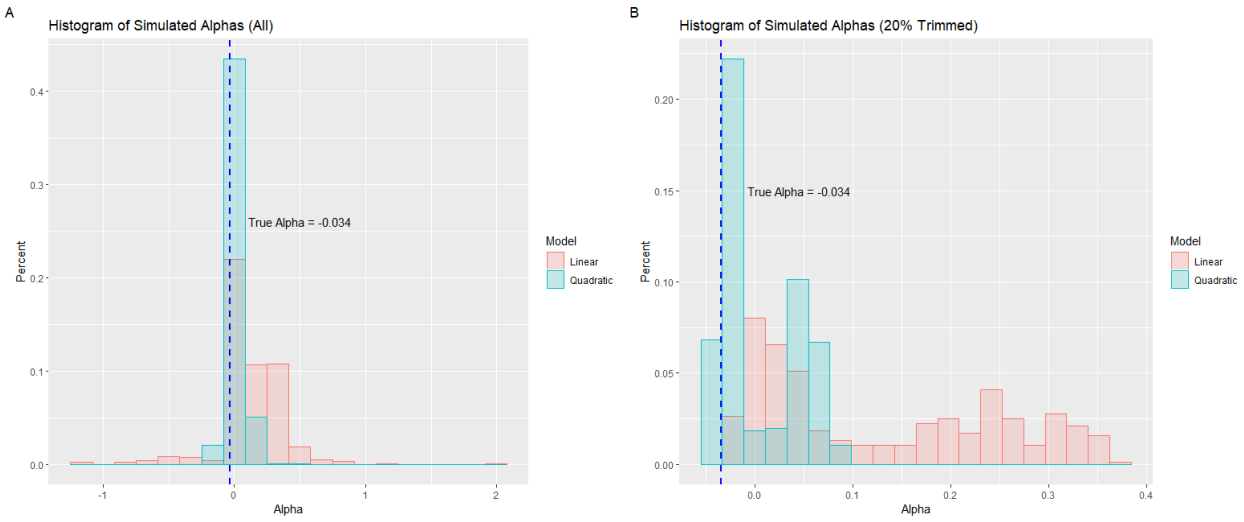

Figure 4 Histogram of Alphas from Simulation 2

Trimming the simulated alphas only offered a marginal improvement in mean bias here (Table 6), with the median faring better, though they are still very biased. In these symptom interference simulations, we can see that misspecifying the longitudinal submodel clearly causes our estimates of the association parameter to be biased in the positive direction.

Table 6 Trimmed and Median Alphas from Simulation 2

	Truth	**Bias**	**% Bias**	**CP**	**MSE**	**Bias**	**% Bias**	**CP**	**MSE**
		Quadratic				**Linear**			
20% Trimmed	-0.034	0.039	-115.3	48.2	0.003	0.162	-474.9	48.6	0.041
Median		0.019	-55.5			0.114	-333.8		

5.0 Discussion

In this thesis we used simulations to examine the effects of misspecifying the longitudinal and survival submodels in joint modeling, using scenarios based on fitting these models to PROs and progression-free survival data in GBM patients. Through applying joint models to the RTOG 0825 data, we confirmed the Gilbert (2014) finding that Bevacizumab patients had a greater deterioration in time in symptom severity and interference scores than placebo patients. Though our models, which were only conducted on a subset of the RTOG 0825 patients, found that the treatment effect was not significantly associated with progression-free survival, we did find that the longitudinal symptom outcomes were significantly associated with progression-free survival, even if the association was very small.

In our simulation based on symptom severity, we found that misspecifying the underlying survival submodel only had a minor impact on estimates of the association parameter in joint models. In fact, the misspecified Weibull submodel appeared to perform even better than the correctly specified exponential submodel, requiring further investigation. In the simulation based on symptom interference, where we misspecified the longitudinal trajectory, we found that fitting a linear trajectory to an assumed quadratic trajectory has a large impact on parameter estimation, especially with respect to estimates of the association parameter, where we see a tendency to overestimate the association parameter. However, we also note that because of the very small association parameters, a small amount of bias could reflect a very large percent bias.

Many issues arose in our analysis and simulation of joint models. We only evaluated the exponential and Weibull underlying survival distributions, which are closely related. Other specifications of the survival submodel could be investigated, including using semi-parametric

over parametric baseline hazards or the rate of change over the current value structure. Some authors suggest modeling the baseline hazard function with splines, especially if there are turning points in the underlying hazard function, something simpler parametric models cannot capture[20]. The RTOG 0825 clinical dataset showed that the survival curves do not meet the proportional hazards assumption, which could be affecting the estimation and simulation of these parameters. Other formulations of the longitudinal submodel should be investigated. Splines have been suggested as a better alternative to polynomial modeling so we can reduce issues where the assumed longitudinal trajectory has a large rate of change as we see in our models. Our simulated longitudinal data would sometimes fall outside the range of the data it was based on, since the MDASI scores are discrete 0-10 scores, which may have influenced our simulated joint models. There was a higher rate of administrative censoring compared to the original data, but when we raised the time limit on administrative censoring, we saw an increased amount of errors in our simulations. This suggests that estimation of joint models can be problematic when there is a large gap between the period of time that longitudinal data is collected and survival events occur. However, this is a common feature of many datasets of this nature and could be an area of further research.

5.1 Conclusion

The results from this thesis show that joint modeling, while a useful tool for incorporating longitudinal and survival analysis together, can be affected by various elements of the data and modeling process. To this end, under our specific formulation of joint models, the estimation of the association parameter, alpha, our primary parameter of interest, seems to be relatively robust

to the misspecification of the underlying survival distribution but not of the true form of the longitudinal trajectory. However, this is difficult to fully evaluate when even our correctly specified simulations gave biased estimates of alpha and our simulated models produced a not insignificant number of extreme estimates of alpha. We find that joint models are effective at evaluating the association between longitudinal and time-to-event outcomes but would advise caution in choosing how to specify the submodels. Though we did not explore this aspect, including additional significant covariates may also improve the estimation of the parameter of interest. We suggest that further research into scenarios when the structure of the data is not ideal and could benefit from more complex submodel specifications would improve the practical application of joint modeling.

Appendix Simulation Code

```
library(JM)
library(plyr)
library(dplyr)
library(mvtnorm)

#####################################
###### Data Description

#### Dataset
#   mdasi: longitudinal dataset
## Variables
#   CN: subject identifier
#   RXf: treatment group assignment
#   Time_Years: time of assessment, in the scale of years
#   SYMPTOM: symptom severity score
#   INF: symptom interference score

#### Dataset
#   surv_m: survival dataset
## Variables
#   PFS: indicator of progression-free survival
#   PFSYears: time of event or censoring, in the scale of years

#####################################
###### Analysis of RTOG 0825 data

## establishing the parameteric regression model
survExp <- survreg(Surv(PFSYears, PFS) ~ RXf, dist = "exponential", data =
surv_m, x = TRUE)

## Mixed Model: square root of symptom severity score
lmeSym <- lme(sqrt(SYMPTOM) ~ RXf*Time_Years + I(Time_Years^2) +
I(Time_Years^3), random = ~ Time_Years | CN, data = mdasi, na.action =
na.omit)

## Mixed Model: square root of symptom interference score
lmeInf <- lme(sqrt(INF) ~ RXf*Time_Years + I(Time_Years^2), random = ~
Time_Years | CN, data = mdasi, na.action = na.omit)

## Joint Model: Symptom Severity Score with Progression-free Survival,
assuming underlying exponential surival distribution
jointFit <- jointModel(lmeSym, survExp, timeVar = "Time_Years", method =
"weibull-PH-aGH", scaleWB = 1)


## Extracting parameters from the joint model to use in simulation
fittedVCV <- jointFit$coefficients$D
longCoef <- jointFit$coefficients$betas
MeasureErr <- jointFit$coefficients$sigma
survCoef <- jointFit$coefficients$gammas[-1]
```

```
assocParm <- jointFit$coefficients$alpha
weibScale <- exp(jointFit$coefficients$gammas[1])
weibShape <- 1/jointFit$coefficients$sigma.t


####################################
###### Simulation Code: using symptom severity joint model as the basis

# Setting number of simulations
NSim <- 500;

# Creating empty lists to store the simulated data and joint models
ModelStorage <- vector("list", NSim)
DataStorage <- vector("list", NSim)
SurvStorage <- vector("list", NSim)
convcodes <- rep(NA, NSim);

for(SimCount in 1:NSim)
{
  ## Simulating random intercept and slopes for N people
  Nsub <- 200;

  ##########################################################
  ### Simulating longitudinal data
  NGroup <- 2; # establishing number of treatment groups
  poly <- 3; # setting order of polynomial time

  ## Specifying true fixed parameters, using parameters extracted above
  b0 <- longCoef[1];
  bTime <- longCoef[3:5]; # Set population time polynomial parameters
  bGroup <- longCoef[c(2,6)]; # Main Effect and interaction

  ## creating the vector of time points for which we simulate data
  tt <- seq(0, 1, by = 0.2)
  Nobs <- length(tt)

  timepolytemp <- matrix(rep(NA, Nobs*poly), nrow = poly, ncol = Nobs)
  for(i in 1:poly)
  {
    timepolytemp[i,] <- tt^i;
  }
  zeroplaceholder <- matrix(rep(0, Nobs*poly), nrow = poly, ncol = Nobs)

  timepoly <- rbind(timepolytemp, zeroplaceholder)

  ## Randomly generating group assignment
  Group <- as.matrix(sample(c(0,1), Nsub, replace = TRUE), ncol = 1)

  ## Specifying and simulating from a multivariate normal distribution
  vcv <- matrix(as.numeric(fittedVCV), nrow = 2)

  mu <- matrix(rep(0,2),ncol = 1)
  U <- as.matrix(rmvnorm(Nsub, mean = mu, sigma = vcv), ncol = 2)

  # Create subject-wise coefficients
  beta0<-b0+U[,1]
  beta1<-bTime[1]+U[,2]
```

```
  beta2<-bTime[2]
  beta3<-bTime[3]
  betaGroup <- bGroup[1]
  betaGroupTime <- bGroup[2]

  s1 <- MeasureErr; # standard deviation of measurement error
  tempE <- rnorm(Nsub*Nobs, mean = 0, sd = s1)
  Eij <- matrix(tempE, nrow = Nsub, byrow = TRUE) # matrix of measurement
error

  # Calculating and constructing the matrix of simulated longitudinal
outcomes
  Y<-matrix(rep(NA,Nsub*Nobs),nrow=Nsub)
  for (i in 1:Nsub) {
    for (j in 1:Nobs) {
      Y[i,j] <- (
beta0[i]+beta1[i]*timepoly[1,j]+beta2*timepoly[2,j]+beta3*timepoly[3,j] +
                    betaGroup*Group[i] +
(betaGroupTime*Group[i]*timepoly[1,j]) + Eij[i,j] )
    }
  }

  ## Adding random censoring
  rcens <- matrix(runif(Nsub*Nobs),nrow=Nsub)
  Y[which(rcens<=0.05)] <- NA


  ##########################################################
  ### Simulating Survival Data  (see section 3.3)

  # Setting parameters based on joint model of real data
  alpha <- assocParm
  Hscale <- weibScale # lambda
  Hshape <- weibShape  # gamma, = 1 if exponential
  ZGroup <- survCoef[1]

  # Generating U from uniform distribution
  UVec <- runif(Nsub*2);
  UMat <- matrix(UVec, ncol = 2)
  logU <- -log(UMat)

  # Creating empty dataframe to hold simulated survival dataset
  tempEventData <- as.data.frame(matrix(rep(NA, Nsub*4), ncol = 4))
  names(tempEventData) <- c("TrueEventTime", "TrueCensTime", "ObsTime",
"ObsEvent")

  # Numerical integration and approximation to solve for T that optimizes
equation (10) from section 3.3
  #  Simulates true survival and censoring times
  for(i in 1:Nsub){
    testfunc <- function(t)
    {
      (t^(Hshape-1))*Hshape*Hscale*exp(ZGroup*Group[i])*
        exp(alpha*(beta0[i] + beta1[i]*t + beta2*t^2 + beta3*t^3 +
betaGroup*Group[i] + betaGroupTime*t*Group[i]))
    }
```

```
    Tvec <- seq(0.001, 3.05, by = 0.001)
    NumInt <- rmutil::int(testfunc, a = 0, b = Tvec)
    tempEventData$TrueEventTime[i] <- Tvec[which.min(abs(NumInt -
logU[i,1]))]
    tempEventData$TrueCensTime[i] <- Tvec[which.min(abs(NumInt - logU[i,2]))]

  }

  # Using T_i and C_i to find V_i and delta_i
  tempEventData$ObsTime <- apply(tempEventData[,1:2], 1, FUN = min)
  tempEventData$ObsEvent <- as.numeric(tempEventData[,3] ==
tempEventData[,1])

    SurvData <- tempEventData[,3:4]
  names(SurvData) <- c("SurvTime", "Event")

  # Set administrative censoring and random censoring
  MaxTime <- 3 # Administrative censoring + reigning in very high survival
times
  SurvData$Marker <- 0
  SurvData$Marker[which(SurvData[,1] > MaxTime)] <- 1
  SurvData$SurvTime[which(SurvData$Marker == 1)] <- MaxTime
  SurvData$Event[which(SurvData$Marker == 1)] <- 0
  SurvData <- SurvData[,-3]


  ########################################################
  ### Data Cleanup, consolidating into readable simulated longitudinal and
survival datasets

  # Subject IDs
  id <- seq(1, Nsub, 1)

  Y_wide <- as.data.frame(cbind(id, Y))
  Y_long <- reshape(Y_wide, direction = "long", varying =
list(names(Y_wide)[2:(Nobs+1)]),
                    v.names = "Y", idvar = "id", timevar = "Time", times =
tt)
  row.names(Y_wide) <- c(1:nrow(Y_wide))

  Y_long <- Y_long[order(Y_long$id, Y_long$Time),]
  row.names(Y_long) <- c(1:nrow(Y_long))

  SurvDataFull <- as.data.frame(cbind(id, Group, SurvData))

  tempJoint <- merge(SurvDataFull, Y_long, by = "id")

  JointData <- tempJoint[-which(tempJoint$Time > tempJoint$SurvTime),]
  JointData <- JointData[-which(is.na(JointData$Y)),] ## removing NAs to
avoid errors

  DataStorage[[SimCount]] <- JointData

  SurvFinal <- aggregate(JointData[,-1], by = list(id=JointData$id),head,1)
  SurvStorage[[SimCount]] <- SurvFinal

  #### Dataset Description
```

```
  ## Variables
  #   id: subject identifier
  #   Group: treatment group assignment
  #   Time: time of longitudinal assessment, in the scale of years
  #   Y: simulated longitudinal outcome
  #   Event: indicator of event or censoring
  #   SurvTime: time of event or censoring, in the scale of years

  ##########################################################
  ### Joint Models

  myLME <- NULL; mySurv <- NULL; myJM <- NULL; # safety measure to ensure old
models aren't reused if a component fails

  # Code to continue running when running into an error
  tryCatch({
    # Mixed model
    myLME <- lme(Y ~ factor(Group)*Time + I(Time^2) + I(Time^3), random = ~
Time | id, data = JointData, na.action = na.omit)
    # Survival Model
    mySurv <- survreg(Surv(SurvTime, Event) ~ factor(Group), dist =
"exponential", data = SurvFinal, x = TRUE)
    # Joint Model
    myJM <- jointModel(myLME, mySurv, timeVar = "Time", method = "weibull-PH-
aGH", scaleWB = 1)

    # Saving joint model output
    ModelStorage[[SimCount]] <- myJM
  }, error = function(e){paste("Error"); ModelStorage[[SimCount]] <- NA})

}
```

Bibliography

1. Vleeschouwer, Steven D. *Glioblastoma*. Brisbane, Australia: Codon Publications, 2017.

2. Brem, Steven, and Kalil G. Abdullah. *Glioblastoma*. Philadelphia, PA: Elsevier, 2017.

3. Hanif, Farina et al. "Glioblastoma Multiforme: A Review of its Epidemiology and Pathogenesis through Clinical Presentation and Treatment." *Asian Pacific journal of cancer prevention : APJCP* vol. 18,1 (2017): 3-9.

4. Gilbert, Mark R et al. "A randomized trial of bevacizumab for newly diagnosed glioblastoma." *The New England journal of medicine* vol. 370,8 (2014): 699-708.

5. Mercieca-Bebber, Rebecca et al. "The importance of patient-reported outcomes in clinical trials and strategies for future optimization." *Patient related outcome measures* vol. 9 (2018): 353-367.

6. Meyers, Christina A et al. "Cognitive function as a predictor of survival in patients with recurrent malignant glioma." *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* vol. 18,3 (2000): 646-650.

7. Rizopoulos, Dimitris. *Joint models for longitudinal and time-to-event data : with applications in R*. Boca Raton: CRC Press, 2012.

8. Tsiatis, Anastasios A., and Davidian, Marie. "An overview of joint modeling of longitudinal and time-to-event data." *Statistica Sinica* vol. 14, (2004): 793-818.

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

10. Brown, Elizabeth R et al. "A flexible B-spline model for multiple longitudinal biomarkers and survival. *Biometrics* vol. 61,1 (2005): 64-73.

11. Henderson, Robin et al. "Joint modelling of longitudinal measurements and event time data." *Biostatistics* vol. 1,4 (2000): 465-480.

12. Gail, Mitchel H et al. "Biased estimates of treatment effect in randomized experiments with nonlinear regressions and omitted covariates." *Biometrika* vol. 71,3 (1984): 431-444

13. Struthers, Cyntha A, and Kalbfleisch, John D. "Misspecified Proportional Hazards Models." *Biometrika* vol. 73, 2 (1986): 363-369

14. Li, Yi-Hwei et al. "Effects of model misspecification in estimating covariate effects in survival analysis for small sample sizes." *Computational Statistics & Data Analysis* vol. 22, 2 (1996): 177-192.

15. Kwong, GPS, and Hutton, JL. “Choice of parametric models in survival analysis: applications to monotherapy for epilepsy and cerebral palsy.” *Journal of the Royal Statistical Society. Series C (Applied Statistics)* vol. 52, 2 (2003): 153-168.

16. Nardi, Alessandar, and Schemper, Michael. “Comparing Cox and parametric models in clinical studies.” *Statistics in Medicine* vol. 22 (2003): 3597-3610.

17. Hutton, Jane L, and Monaghan, PF. “Choice of Parametric Accelerated Life and Proportional Hazards Models for Survival Data: Asymptotic Results.” *Lifetime Data Analysis* vol. 8 (2002): 375-393.

18. Bender, Ralf et al. “Generating survival times to simulate Cox proportional hazards models.” *Statistics in Medicine* vol. 24 (2005): 1713-1723.

19. Crowther, Michael J, and Lambert, Paul C. “Simulating biologically plausible complex survival data.” *Statistics in Medicine* vol. 32, 23 (2013): 4118-4134.

20. Crowther, Michael J et al. “Joint modelling of longitudinal and survival data: incorporating delayed entry and an assessment of model misspecification.” *Statistics in Medicine* vol. 35, 7 (2016): 1193-1209.