TOPICS IN JOINT MODELING OF LONGITUDINAL BIOMARKER, QUALITY OF LIFETIME, AND SURVIVAL DATA

by

Xinxin Dong

MS, University of Pittsburgh, 2009

BMed, Beijing University, China, 2007

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented

by

Xinxin Dong

It was defended on

July 8th 2013

and approved by

Abdus S. Wahed, Ph.D., Associate Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Lan Kong, Ph.D., Associate Professor, Division of Biostatistics and Bioinformatics, Penn State University

Carol K. Redmond, ScD, Distinguished Service Professor Emerita, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Joyce Chung-Chou Ho Chang, Ph.D., Associate Professor, Department of Medicine, Biostatistics, and Clinical and Translational Science, University of Pittsburgh

Douglas P. Landsittel, Ph.D., Professor, Department of Medicine, Biostatistics, and Clinical and Translational Science, University of Pittsburgh

Dissertation Director: Abdus S. Wahed, Ph.D., Associate Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

TOPICS IN JOINT MODELING OF LONGITUDINAL BIOMARKER, QUALITY OF LIFETIME, AND SURVIVAL DATA

Xinxin Dong, PhD

University of Pittsburgh, 2013

ABSTRACT: Joint modeling techniques have been developed for analyzing correlated longitudinal and survival data in many studies. It provides consistent and efficient estimates of the parameters even when the longitudinal covariate is measured infrequently and with measurement error. This work focuses on the use of joint modeling to solve two different statistical problems. The first one is about analyzing censored biomarker measurements and survival data under a case-cohort design. The goal is to study how the biomarker level changes over time, and the relationship between longitudinal biomarker measurements and the event time. We suggest a modified likelihood-based approach to adjust the possible bias introduced by the censoring due to detection limits and the measurement error in biomarker measurements. The second topic is about drawing inference for mean quality adjusted lifetime data. We consider continuous health experience and define the quality function with repeatedly measured quality of life score. We propose a consistent and asymptotic normal estimate for the mean quality adjusted lifetime and derive its asymptotic variance. The performances of the proposed methods have been demonstrated in simulation studies and through real data examples.

Public Health Significance: A case-cohort study is a cost-effective design that is used in many large epidemiological studies. The method proposed in the first part of this dissertation increases the efficiency of the parameter estimation under case-cohort designs, which can lead to a considerable reduction in cost and effort. The evaluation of health benefits is a major public health challenge. The second part of this dissertation presents the estimate of

mean quality adjusted lifetime, which can serve as a measurement of health benefits in the comparison of treatments or public health strategies.

Keywords: Joint analysis, Case-cohort, Longitudinal biomarker, Limit of detection (LOD), Mixed effects model, Accelerated failure time model, Quality of life, Survival analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

It is difficult to overstate my sincere gratitude to my thesis advisor, Dr. Abdus Wahed. I greatly thank him for his inspiration, patience, and encouragement. I enjoyed the meetings with him and was always benefited from his insightful comments and suggestions. I also own my deepest gratitude to my former advisor, Dr. Lan Kong. She made herself available whenever I needed her and extended abundant time and effect to help me write the manuscript and guide me with patience. I was fortunate to have Dr. Carol Redmond as my master advisor and committee member for the PhD dissertation. She taught me not only professional knowledge, but also how to be a good researcher. I am indebted to Dr. Joyce Chang in many ways. Being a teaching assistant for her class provided me with a valuable opportunity to learn the presentation skills from the best, and the Graduate Student Researcher (GSR) position she recommended helped me gain extremely precious research experience. I would like to acknowledge Dr. Douglas Landsittel for carefully reading my thesis, and giving me helpful suggestions about the practical use of the methodologies. His input has improved this dissertation and also inspired ideas for future research direction.

I am grateful to all the faculty members, colleagues and friends in the Department of Biostatistics at the University of Pittsburgh. They have made a significant impact on my professional growth. What I learned from them will be a lifetime treasure.

Finally, much gratitude is owed to my parents and husband for all of their love, support and encouragement.

1.0 INTRODUCTION

In the first part of this dissertation, we describe a joint analysis approach for biomarker and survival data under a case-cohort design. Biomarkers are often measured over time in epidemiological studies and clinical trials for understanding better the mechanism of diseases. In large cohort studies, case-cohort sampling provides a cost effective method to collect expensive biomarker data for revealing the relationship between biomarker trajectories and time to event. However, biomarker measurements are often limited by the sensitivity and precision of a given assay, resulting in data that are censored at detection limits and prone to measurement errors. Additionally, occurrence of an event of interest may proclude biomarkers from being further evaluated. Inappropriate handling of these types of data can lead to biased estimates or errorneous conclusions. Under a classical case cohort design, we propose a modified likelihood-based approach to accommodate these special features of longitudinal biomarker measurements in the accelerated failure time (AFT) models. The maximum likelihood estimators based on the full likelihood function are obtained by the Gaussian quadrature method. We evaluate the performance of our case-cohort estimator and compare its relative efficiency to the full cohort estimator through simulation studies. The proposed method is further illustrated using the data from a biomarker study of sepsis among patients with community acquired pneumonia.

The second part of this dissertation focuses on quality adjusted lifetime, which has become important measure for treatment evaluation in clinical trials or observational studies to quantify health benefits. Zhao and Tsiatis (2000) [63] derived a class of estimators and their asymptotic variances for mean quality adjusted lifetime for right-censored survival data. In their consideration, the quality function was assumed to be a known piecewise constant function where the quality remains constant within the same state. Consequently, they defined

the quality adjusted lifetime as a weighted sum (integral) of the times spent at each health state until the event of interest or censoring. However, in many applications quality of life is repeatedly measured using instruments over time and hence the actual quality function is unknown and ususally estimated by fitting some non-linear model over time. We consider inference on mean quality adjusted lifetime (QAL) when the quality function is estimated from the data. Specifically, we propose an estimator of the mean QAL and establish its asymptotic properties. The proposed estimator is shown to be consistent and asymptotically normal. The estimator of its asymptotic variance is also derived. The practical performance of the proposed estimator is illustrated via a simulation study.

2.0 ACCELERATED FAILURE TIME MODEL FOR CASE-COHORT DESIGN WITH LONGITUDINAL COVARIATES SUBJECT TO MEASUREMENT ERROR AND DETECTION LIMITS

2.1 INTRODUCTION

Longitudinal measurements collected along with time-to-event information have become an important data component in many epidemiological and clinical studies. Common objectives of such studies are to examine how the longitudinal variables vary over time, and how they correlate with the event of interest. For instance, a panel of biological markers is often measured over time to understand better the mechanism of a disease and aid in the development of effective treatments. In vaccine studies, immune responses are evaluated repeatedly to see if they can serve as surrogate markers for the study endpoint which takes a long time to be observed. Marker information for environmental exposure is assembled over time to study how the history of being exposed to certain chemical materials is associated with the risk of an occupational disease. Our motivating example, Genetic and Inflammatory Markers of Sepsis (Kellum et al., 2007) [25] study, focused on understanding the natural history and development of sepsis, and identifying the biomarkers indicating the risk for severe sepsis, multiple organ failure, and death. A set of inflammatory and coagulation markers from blood samples were evaluated repeatedly during the course of hospitalization, and one of the primary outcomes was 90-day mortality since enrollment.

In some large cohort studies, it is prohibitive to analyze multiple biomarkers over time for each individual although the biological samples can be collected for all subjects in the cohort. A cost effective solution is the case-cohort design (Prentice, 1986) [40], where a random sample, namely subcohort, is selected from a well-defined cohort at the beginning of

the study. Covariate data of interest are collected for all subjects in this subcohort and those who develop the event of interest outside the subcohort. Compared with full-cohort study, a case-cohort design leads to significant reductions in cost and effort, especially when the event of interest is rare and/or exposures are expensive to ascertain. It is also flexible for analyzing multiple outcomes of interest by using the subcohort as the common comparison group. The Atherosclerosis Risk in Communities study (The ARIC Investigators, 1989) [49] used a case-cohort design to investigate certain plasma/genetic markers as risk factors of coronary heart disease (CHD) and stroke. The Busselton Health study [10], a large Australian cohort study, applied case-cohort sampling to study the association between serum markers and CHD and stroke events to reduce costs and preserve stored serum.

Time-to-event data under a case-cohort design have been extensively studied using the proportional hazards (PH) model (Cox, 1972) [9]. A pseudo-likelihood approach for inference from the PH model has been proposed by Prentice (1986) [40], and further investigated by Self and Prentice (1988) [45], and Lin and Ying (1993) [32]. Variance estimators in the PH model have been presented by Wacholder et al. (1989) [56] and Barlow (1994) [3]. Computational issues have been described and addressed by Therneau and Li (1999) [50], Langholz and Jiao (2007) [31]. Improved estimators have been derived by Chen and Lo (1999) [6], Borgan et al. (2000) [4], Chen (2001a) [7], and Samuelsen et al. (2007) [44]. Sorensen and Anderson (2000) [47] considered a competing risk model for case-cohort data. Non-proportional hazard models have also been studied under a case-cohort design, for example, Kulich and Lin (2000) [30], Sun et al. (2004) [48], and Ma (2007) [37] proposed an additive hazards model. Chen (2001b) [5], Kong et al. (2004) [27], Lu and Tsiatis (2006) [35] considered semiparametric transfromation models. Nan et al. (2006) [38], Kong and Cai (2009) [26] studied the case-cohort data under accelerated failure time models. Lu and Shih (2006) [35], and Zhang et al. (2011) [61] proposed methods to accommodate clustered survival data. Kang and Cai (2009) [28] considered multiple events under a case-cohort design.

Longitudinal variables can be incorporated into survival models as time-dependent covariates, if they were measured without error and complete histories were known. Challenges arise, however, because biomarker information are usually collected intermittently,

interrupted by the event time, and often subject to measurement error. Furthermore, due to the sensitivity and precision of a given assay, biomarker measurements are censored by limit of detection (LOD) when the concentrations are higher or lower than certain thresholds. Given this nature of longitudinal data, joint modeling with a modified likelihood function accounting for censoring due to LOD could be an ideal approach to examine if the exposure trajectory (time-dependent covariate) influences the survival outcome. Joint analysis of time-to-event and longitudinal data has been considered by Faucett and Thomas (1996) [12], Hogan and Laird (1997a, 1997b, 1998) [20] [21], Wulfsohn and Tsiatis (1997) [58], Henderson et al. (2000) [19], Hsieh et al. (2006) [22] and Vonesh et al. (2006) [55], among many others. Under a nested case-control design, Tseng and Liu (2009) [51] linked a mixed-effect model with a Cox model by shared latent random effects. Inference is drawn based on the maximum likelihood method and is applicable to a case-cohort design. Most of the joint modeling papers have focused on the use of Cox proportional hazards model to study the association between repeated measures and survival information. However, there are two main limitations of the Cox model: 1) the interpretation of covariate effect is not directly on the survival time, but through the hazard function which may not be intuitive to the clinical investigators; and 2) proportionality of hazards may not hold true in some applications.

An accelerated failure time (AFT) model (Cox, 1972) is an attractive alternative due to its appealing interpretation and robustness to non-proportionality. It can be considered as a log linear model that links the log transformed survival time to a set of covariates by a linear regression model. The role of covariates can be interpreted as either accelerating or decelerating the time to event. Tseng et al. (2005) [52] proposed a joint model of accelerated failure time and longitudinal data for a full cohort study. We considered applying joint modeling approach to case-cohort studies by using a linear mixed effects model to describe the trajectory of longitudinal covariates, and an accelerated failure time model to investigate the relationship between survival outcomes and longitudinal covariates subject to detection limits. There are several advantages of our proposed method: 1) the model we use to describe the covariate process allows the variation of population means over time, the unobserved heterogeneity among subjects and possible additive measurement errors; 2) we efficiently use all available information by incorporating the survival and covariate information of subjects

who have longitudinal measurements missing by design; and 3) the likelihood we propose accounts for longitudinal biomarkers data censored due to detection limits. Our estimation procedure can be easily implemented when the baseline hazard function is specified. We apply the Gaussian quadrature method through NLMIXED procedure in SAS to obtain the maximum likelihood estimates for parameters in both longitudinal and survival models simultaneously.

In Section 2.2 we introduce the details of our model specifications. Section 2.3 presents the joint likelihood and estimation procedure. Section 2.4 describes the results from simulation studies. Section 2.5 illustrates the method using the biomarker data from the Genetic and Inflammatory Markers of Sepsis (GenIMS) study. Finally, Section 2.6 provides discussion and concluding remarks.

2.2 NOTATION AND MODEL SPECIFICATIONS

Suppose that the full study cohort includes N subjects. For individual i, let T_i, C_i, and X_i be the event time, the potential right censoring time, and the baseline covariate vector, respectively. Due to censoring, we observe (Y_i, Δ_i, X_i), where the time on study $Y_i = min(T_i, C_i)$ and the event indicator $\triangle_i = 1$, if $T_i \leq C_i$, and 0 otherwise. At the beginning of the study, a random sample, namely subcohort, is selected from the entire cohort. The case-cohort subjects comprise this subcohort and all other cases. Let $R_i = 1$ if the ith individual is in the case-cohort; and $R_i = 0$, otherwise. For individuals with $R_i = 1$, the longitudinal covariate process, $Z_i(t)$, is measured at $t_i = \{t_{ij}, j = 1, \ldots, m_i\}$, for $t_{im_i} \leq Y_i$, with measurement error $\epsilon_i = \{\epsilon_{ij}, j = 1, \ldots, m_i\}$.

Let $W_{ij} = Z_i(t_{ij}) + \epsilon_{ij}$ be the observed measurement of covariate Z at time t_{ij} for the ith subject. We postulate a linear mixed effects model for the observed longitudinal covariate W_{ij} as

$$W_{ij} = b_i^T G_{ij} + \eta^T X_{ij}^* + \epsilon_{ij}, \tag{2.1}$$

where G_{ij} is the subject specific design vector for the random effect b_i, X^*_{ij} is the design vector for the fixed effects which might contain individual baseline covariates and the measurement times t_{ij}, and ϵ_{ij}'s are the measurement errors, assumed to be independently normally distributed with mean zero and variance σ^2. Generally, the random effect b_i's are also assumed to be normally distributed with mean zero and covariance matrix Σ and they are independent of the measurement errors ϵ_{ij}'s. Model (3.3) is the simple linear mixed effects model that accounts for within subjects correlation through subject-specific random effects, and the fixed parameters of this model allow assessment of change in the longitudinal biomarker process over time. More complicated models, e.g. models that incorporate baseline covariate and time interaction, can be considered in a similar fashion.

Due to the sensitivity of a given instrument, W_{ij} is measurable within a detectable range and can be censored by the lower and upper limits of detection, denoted by τ_l and τ_u, respectively. Let $d_i = \{d_{ij}, j = 1, \ldots, m_i\}$ be the censoring indicator vector for longitudinal measurements, then $d_{ij} = 0$ if $\tau_l \leq W_{ij} \leq \tau_u$; 1 if $W_{ij} < \tau_l$; and 2 if $W_{ij} > \tau_u$. Therefore, the observed data for subject i can be written in the form of $(Y_i, \Delta_i, X_i, R_i, R_i\{W_{ij}^{I(d_{ij}=0)}\tau_l^{I(d_{ij}=1)}\tau_u^{I(d_{ij}=2)}, j = 1, 2, \ldots, m_i\}, R_i t_i)$, $i = 1, 2, ..., n$.

To evaluate the risk factors of time to event outcome, accelerated failure time (AFT) model provides an important alternative to Cox regression model because it models the failure time directly and allows non-proportionality of hazards. If the covariates are time-invariant, AFT model relates the natural logarithm of event time to a set of covariates by a linear regression model,

$$log(T_i) = -\beta' X_i + \epsilon_i, \tag{2.2}$$

where β is an unknown vector of regression coefficients, β' denotes the transpose of β, and ϵ_i's are independently and identically distributed random errors. Let T_0 be the event time variable at $X = 0$, then $T_i = T_0 \exp(-\beta' X_i)$ where $T_0 = \exp(\epsilon)$. Suppose $S_0(t)$ is the survival function of T_0, which is also referred to as baseline survival function. Model (2.2) implies that the survival function of T at given covariate vector X is $S(t|X) = S_0(t \exp\{\beta' X\})$. Thus, the covariate in AFT model has a multiplicative effect on the failure time rather than on the hazard function as in the Cox regression models.

To evaluate the effect of the longitudinal process on the event time, we use the extension of an AFT model for time-dependent covariates suggested by Cox and Oakes (1984). Define the "covariate history" up to t as $\bar{H}_i(t, b_i) = \{H_i(s, b_i), s < t\}$ where $H_i(s, b_i) = \{X_i, Z_i(s)\}$. Note that, even though we are referring to $H_i(.)$ as covariate history, it actually includes the baseline covariates as well as the random effects through the conditional mean of the longitudinal process as described in Equation (3.3). Thus $H_i(.)$ is not completely observable. The time-dependent AFT model is represented by

$$T_0 = \int_0^{T_i} \exp\{\beta' H_i(s, b_i)\} ds, \tag{2.3}$$

with $T_0 = \exp(\epsilon)$ as defined before. Therefore, the hazard function conditional on $\bar{H}_i(t, b_i)$ can be expressed by

$$\lambda(t|\bar{H}_i(t, b_i)) = \lambda_0 \left(\int_0^t \exp\{\beta' H_i(s, b_i)\} ds \right) \exp\{\beta' H_i(t, b_i)\}, \tag{2.4}$$

where λ_0 is the hazard function of T_0, and is usually referred to as baseline hazard function.

In this paper, we focus on the case when the explicit form of λ_0 is known. In other words, the distribution of error term ϵ is known. As indicated in Models (2.3) and (2.4), the entire history of the longitudinal process influences the hazard for subject i at time t, while in the time-dependent Cox model, the hazard function depends on the covariate history only through the current covariate value.

2.3 INFERENCE PROCEDURE UNDER A CASE-COHORT DESIGN

Under a case-cohort design, the longitudinal information of controls outside the subcohort are missing by design. It is natural to apply the statistical methods for missing data problem to case-cohort data. Adopting the framework presented by Zeng et al. (2006) [60] and Tseng and Liu (2009) [51], we assume that the longitudinal data is missing at random (MAR) and noninformative (Little and Rubin, 2002) [33], because the probability of missing only depends on the observed event histories and usually some baseline covariates under case-cohort design.

We further assume that the measurement schedule t_i is independent of baseline covariates X_i, future longitudinal measurements, and random effects b_i. Therefore the statistical inference can be based on the observed data. To handle the censored biomarker measurements due to LODs, we use the method similar to Lyles et al.(2000) [36]. The likelihood function is constructed using the conditional density function $f_w(W_{ij}|b_i)$ for observed measurements and the cumulative distribution function $F_w(\tau_l|b_i)$ or $1-F_w(\tau_u|b_i)$ for censored observations. The observed data likelihood function for parameters of interest, $\phi = (\beta, \eta, \Sigma, \sigma^2)'$, can then be written as

$$L(\phi, \lambda_0) = \prod_{i=1}^{n} l_i(\phi, \lambda_0) \tag{2.5}$$

where

$$l_i(\phi, \lambda_0) = \begin{cases} \int \left\{ \prod_{j=1}^{m_i} f_w(W_{ij}|b_i)^{I(d_{ij}=0)} F_w(\tau_l|b_i)^{I(d_{ij}=1)} (1 - F_w(\tau_u|b_i))^{I(d_{ij}=2)} \right\} & \\ \times f(Y_i, \Delta_i|b_i) f_b(b_i) db_i & \text{if } R_i = 1 \\ \int f(Y_i, \Delta_i|b_i) f_b(b_i) db_i & \text{if } R_i = 0 \end{cases} , \tag{2.6}$$

$f(Y_i, \Delta_i|b_i)$ and $f_b(b_i)$ respectively define the likelihood contribution from the survival component, and the marginal density of the random effect b_i. For our study, we considered the following specific models:

$$\begin{aligned} f_w(w|b_i) &= (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{ -\frac{[w - Z(t_{ij}, b_i)]^2}{2\sigma^2} \right\}, \\ f(y, \delta|b_i) &= \left[\lambda_0 \left(\int_0^y \exp\{\beta' H(s, b_i)\} ds \right) \exp\{\beta' H(y, b_i)\} \right]^{\delta} \\ &\quad \times \exp\left\{ - \int_0^{\int_0^y \exp\{\beta' H(s,b_i)\} ds} \lambda_0(u) du \right\}, \\ f_b(b) &= (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\{-b'\Sigma^{-1} b/2\}, \end{aligned} \tag{2.7}$$

where p is the dimension of random effect b_i. Note that, for simplicity, we have dropped the fixed parameters from the density notation. Traditional case-cohort analysis uses the data from case cohort members only, here we also incorporate the information of event time and any available covariate information for the controls outside the subcohort ($R_i = 0$). With fully specified baseline hazard function λ_0, we have a parametric form of an

accelerated failure time model. Liu and Huang (2008) [34] suggested the use of Gaussian quadrature for estimation of proportional hazards models including fraity terms. We adopt this estimation method and apply the Gaussian quadrature technique to approximate the integrals in Equation (2.5) by weighted sums over predefined abscissas for the random effects.

We set the number of abscissas as five because numerical studies have shown that Gaussian quadrature with five quadrature points provides a good approximation to the true likelihood function, and the logarithm of $\widetilde{L}(\phi, \lambda_0)$ can be maximized by the quasi-Newton method (Liu and Huang, 2008 [34]).

The implementation of this approach is straightforward using Proc NLMIXED procedure of SAS 9.3, because no term is left unspecified in the likelihood function and the random effects are assumed to follow a multivariate normal distribution. However, similar to many other non-linear models, the choice of initial values is critical in terms of the accuracy of the final estimate and the computational time. We apply a two-stage approach to acquire initial values: first, we model longitudinal data with a mixed effects model; then we use the resulting estimates as if they were observed to obtain the estimates for the AFT model based on the likelihood function in (2.7). The estimates from two-stage analysis are used as the initial values for the estimation algorithm. Our estimators are expected to retain all the advantageous asymptotic properties of maximum likelihood estimator (MLE), such as consistency, efficiency and asymptotic normality under certain regularity conditions. The corresponding standard errors can be directly obtained from the inverse, or generalized inverse if singular, of negative Hessian matrix. Therefore, formal inference, e.g. confidence intervals and tests of hypothesis, can be conducted via the Wald method.

2.4 SIMULATION STUDIES

We carry out a number of simulation studies to investigate the performance of the proposed estimation procedure and assess the efficiency of the case-cohort design compared to the full cohort design. We generate longitudinal and survival data from three joint models as follows.

The first joint model (M1) consists of: (i) a mixed effects model with only random

intercept:

$$W_{ij} = l_0 + l_1 t_{ij} + b_{0i} + \epsilon_{ij}, \tag{2.8}$$

where the fixed intercept and slope, (l_0, l_1), represent the overall trend of longitudinal measurements over time; the random intercept, $b_{0i} \sim \mathcal{N}(\mathbf{0}, \sigma_{\mathbf{0}}^{\mathbf{2}})$, reflect the heterogeneity among subjects; and the random errors, $\epsilon_{ij} \sim \mathcal{N}(\mathbf{0}, \sigma^{\mathbf{2}})$ represent the variation within subject and are independent of random effects; (ii) an AFT model, assuming the survival time is related to the longitudinal measurements through the random intercept only:

$$T_0 = T_i \exp\{\beta b_{0i} + \eta X_i\}, \tag{2.9}$$

where T_0 follows a weibull distribution with a shape parameter a and a scale parameter b. The second joint model (M2) consists of: (i) the same mixed effects model as described in M1; (ii) an AFT model, assuming the survival time is related to the true underlying longitudinal process:

$$T_0 \;=\; \int_0^{T_i} \exp\{\beta(l_0 + l_1 s + b_{0i}) + \eta X_i\} ds, \tag{2.10}$$

where T_0 follows an exponential distribution with hazard rate λ_0 and X_i is a vector of baseline covariates that affects the survival time. The third joint model (M3) consists of: (i) a mixed effects model with both random intercept and slope:

$$W_{ij} = l_0 + l_1 t_{ij} + b_{0i} + b_{1i} t_{ij} + \epsilon_{ij}, \tag{2.11}$$

where l_0, l_1, b_{0i} and ϵ_{ij} are the same as described in M1, random slope $b_{1i} \sim \mathcal{N}(\mathbf{0}, \sigma_{\mathbf{1}}^{\mathbf{2}})$ and is independent of b_{0i}; (ii) an AFT model given by

$$T_0 \;=\; \int_0^{T_i} \exp\{\beta(l_0 + l_1 s + b_{0i} + b_{1i} s) + \eta X_i\} ds, \tag{2.12}$$

where T_0 follows an exponential distribution with hazard rate λ_0. Ten measurements are preliminarily scheduled for each subject i at $t_i = (1, 2, ..., 10)$, and no measurement is available after the occurrence of the event of interest or censoring. The baseline covariate X_i is drawn from a binomial distribution with probability 0.5. To allow for various degrees of association between the longitudinal process and event time, we set β to be 0.1 and 0.5. The other parameters are specified as: $a = 2, b = 2, l_0 = 1, l_1 = 2, \eta = 1, \sigma_0^2 = 1$, and

$\sigma^2 = 1$ for M1; $\lambda_0 = 0.002, l_0 = 0.5, l_1 = 10, \eta = 1, \sigma_0^2 = 0.5$, and $\sigma^2 = 1$ for M2; and $\lambda_0 = 0.002, l_0 = 0.5, l_1 = 10, \eta = 1, \sigma_0^2 = 0.25, \sigma_1^2 = 0.25$, and $\sigma^2 = 1$ for M3.

The censoring time, C_i, is generated from exponential distributions with hazard rates chosen appropriately such that the resulting proportions of cases are expected to be about 20% or 10%. We generate 500 full cohort data sets under each simulation with sample sizes of N=600 and 1200. For each data set, we empirically choose detection limits to achieve 10%, 20%, and 40% censoring rates for the longitudinal measurements. We select subcohorts by two different sampling proportions such that the ratios of cases to controls in the resulting case-cohort studies are 1:1 and 1:2, respectively. We obtain the maximum likelihood estimators and their standard errors by the method proposed in section 3.

Tables 2.4, 2.4, and 2.4 summarize the performance of the proposed estimators for β, which is of our primary parameter of interest, under three models, M1, M2 and M3, respectively. The results for 20% censoring rate due to detection limits were similar and hence are not presented here. As indicated in the simulation results, the proposed estimators are approximately unbiased in all cases. The biases of the proposed estimators are all minimal. Furthermore, the means of the estimated standard errors are in good agreement with the empirical standard deviations of the estimates. The coverage rates of 95% confidence intervals match the nominal level. We compute the relative efficiency (RE) of $\hat{\beta}$ as the ratio of mean square errors (MSE) of the full cohort estimator to the case-cohort estimator. As shown in Table 1, for $\beta = 0.5$, under 10% failure rate, the case-cohort estimate of β retains about 70% efficiency of the full cohort estimate when longitudinal covariate information is obtained from only 20% (1:1 case-to-control ratio) of the full cohort subjects. When the effect of longitudinal covariate is weaker ($\beta = 0.1$), the loss of efficiency becomes larger, however at least 60% efficiency is achieved with 30% of the full cohort subjects. As expected, the efficiency of the case-cohort estimator increases with case-cohort sample size and failure rate. Our method still performs well when the censoring proportion of longitudinal covariate measurements is as high as 40%. Similar results are seen in Tables 2.4 and 2.4 for the cases where the survival time is related to the longitudinal covariate through the entire covariate process rather than just the random intercept as in M1.

The biases of the proposed estimators for other parameters in Models (2.9), (2.10) and (2.12) are relatively small (results not presented). The standard deviations of the case-cohort estimates for η are almost the same as full-cohort estimates, indicating that for the covariate that is available for every subject, our case-cohort estimator is as efficient as the full-cohort estimator. Comparing to full-cohort estimates, the efficiency loss is small for the case-cohort estimates of the parameters that specify the baseline survival function. But the case cohort estimates for the regression parameters in the mixed effects model show relatively larger efficiency loss.

Table 1: Simulation results for β under model M1 where T_0 follows a Weibull distribution with 500 replications.

P_c	β	N	Design	20% cases					10% cases				
				Bias($\hat{\beta}$)	MeanSE	SD($\hat{\beta}$)	95%CP	RE	Bias($\hat{\beta}$)	MeanSE	SD($\hat{\beta}$)	95%CP	RE
0.1	0.1	600	Full cohort	0.001	0.048	0.048	0.948	—	0.004	0.073	0.080	0.932	—
			CC 1:2	0.001	0.053	0.055	0.936	0.822	0.008	0.092	0.099	0.932	0.635
			CC 1:1	0.002	0.058	0.059	0.940	0.684	0.003	0.101	0.110	0.924	0.522
		1200	Full cohort	0.004	0.034	0.035	0.946	—	0.003	0.051	0.051	0.952	—
			CC 1:2	0.003	0.038	0.039	0.932	0.824	0.001	0.063	0.063	0.950	0.649
			CC 1:1	0.004	0.041	0.042	0.944	0.691	0.001	0.070	0.074	0.940	0.526
	0.5	600	Full cohort	0.002	0.058	0.057	0.956	—	0.003	0.077	0.081	0.952	—
			CC 1:2	0.002	0.062	0.063	0.944	0.878	<0.001	0.088	0.090	0.946	0.779
			CC 1:1	0.003	0.065	0.066	0.948	0.801	0.001	0.093	0.097	0.938	0.697
		1200	Full cohort	<0.001	0.041	0.039	0.952	—	0.002	0.055	0.055	0.946	—
			CC 1:2	<0.001	0.043	0.042	0.962	0.883	0.004	0.062	0.065	0.952	0.786
			CC 1:1	<0.001	0.045	0.044	0.960	0.806	0.005	0.065	0.066	0.942	0.711
0.4	0.1	600	Full cohort	0.001	0.049	0.049	0.952	—	0.005	0.075	0.082	0.928	—
			CC 1:2	0.001	0.054	0.056	0.934	0.823	0.009	0.094	0.102	0.936	0.633
			CC 1:1	0.003	0.059	0.061	0.942	0.682	0.003	0.103	0.114	0.918	0.525
		1200	Full cohort	0.003	0.035	0.036	0.942	—	0.002	0.052	0.052	0.952	—
			CC 1:2	0.003	0.038	0.040	0.946	0.819	0.001	0.064	0.065	0.948	0.650
			CC 1:1	0.004	0.042	0.043	0.948	0.685	<0.001	0.071	0.076	0.946	0.525
	0.5	600	Full cohort	0.001	0.061	0.059	0.952	—	0.004	0.082	0.083	0.956	—
			CC 1:2	0.002	0.066	0.066	0.948	0.862	0.001	0.094	0.094	0.952	0.756
			CC 1:1	0.004	0.070	0.070	0.950	0.769	0.001	0.100	0.102	0.944	0.662
		1200	Full cohort	0.001	0.043	0.042	0.958	—	0.002	0.058	0.059	0.948	—
			CC 1:2	<0.001	0.046	0.045	0.962	0.868	0.005	0.066	0.070	0.954	0.763
			CC 1:1	<0.001	0.049	0.048	0.964	0.777	0.005	0.070	0.072	0.954	0.676

$a = 2, b = 2, l_0 = 1, l_1 = 2, \eta = 1, \sigma_0^2 = 1$, $\sigma^2 = 1$. P_c refers to proportion of longitudinal measurements censored due to detection limit. CC 1:2 and CC 1:1 refer to case-cohort study with case-to-control ratio as 1:2 and 1:1, respectively. RE refers to the relative efficiency of case-cohort estimator, defined by the ratio of mean square errors(MSE) with the full cohort design as a reference.

Table 2: Simulation results for β under model M2 where T_0 follows an Exponential distribution with 500 replications.

P_c	β	N	Design	20% cases					10% cases				
				Bias($\hat{\beta}$)	MeanSE	SD($\hat{\beta}$)	95%CP	RE	Bias($\hat{\beta}$)	MeanSE	SD($\hat{\beta}$)	95%CP	RE
0.1	0.1	600	Full cohort	0.001	0.008	0.008	0.954	—	0.002	0.010	0.011	0.948	—
			CC 1:2	0.002	0.009	0.009	0.958	0.781	0.003	0.013	0.014	0.954	0.620
			CC 1:1	0.002	0.010	0.011	0.946	0.636	0.003	0.015	0.015	0.964	0.491
		1200	Full cohort	0.001	0.006	0.006	0.942	—	0.001	0.007	0.007	0.944	—
			CC 1:2	<0.001	0.006	0.006	0.952	0.801	0.002	0.009	0.009	0.960	0.630
			CC 1:1	0.001	0.007	0.007	0.958	0.644	0.002	0.010	0.010	0.954	0.508
	0.5	600	Full cohort	0.007	0.040	0.040	0.944	—	0.013	0.048	0.047	0.960	—
			CC 1:2	0.007	0.045	0.046	0.950	0.813	0.016	0.061	0.057	0.980	0.634
			CC 1:1	0.009	0.050	0.049	0.964	0.650	0.016	0.068	0.067	0.966	0.514
		1200	Full cohort	0.002	0.028	0.027	0.948	—	0.007	0.034	0.034	0.960	—
			CC 1:2	0.002	0.031	0.030	0.950	0.809	0.010	0.042	0.043	0.960	0.614
			CC 1:1	0.004	0.035	0.034	0.950	0.644	0.010	0.047	0.047	0.954	0.505
0.4	0.1	600	Full cohort	0.001	0.008	0.008	0.960	—	0.002	0.011	0.011	0.956	—
			CC 1:2	0.002	0.010	0.010	0.956	0.764	0.003	0.014	0.015	0.956	0.587
			CC 1:1	0.002	0.011	0.011	0.940	0.611	0.004	0.016	0.016	0.968	0.462
		1200	Full cohort	0.001	0.006	0.006	0.952	—	0.001	0.007	0.008	0.942	—
			CC 1:2	<0.001	0.007	0.007	0.960	0.783	0.002	0.010	0.010	0.956	0.598
			CC 1:1	0.001	0.008	0.007	0.970	0.622	0.002	0.011	0.011	0.956	0.478
	0.5	600	Full cohort	0.007	0.041	0.041	0.952	—	0.013	0.049	0.049	0.962	—
			CC 1:2	0.006	0.046	0.048	0.944	0.807	0.016	0.064	0.061	0.962	0.609
			CC 1:1	0.008	0.052	0.050	0.966	0.632	0.017	0.072	0.073	0.976	0.479
		1200	Full cohort	0.002	0.029	0.028	0.962	—	0.007	0.034	0.034	0.960	—
			CC 1:2	0.003	0.032	0.032	0.960	0.796	0.011	0.044	0.045	0.950	0.592
			CC 1:1	0.005	0.037	0.036	0.954	0.619	0.010	0.050	0.049	0.958	0.479

$\lambda_0 = 0.002, l_0 = 0.5, l_1 = 10, \eta = 1, \sigma_0^2 = 0.5$, $\sigma^2 = 1$. P_c refers to proportion of longitudinal measurements censored due to detection limit. CC 1:2 and CC 1:1 refer to case-cohort study with case-to-control ratio as 1:2 and 1:1, respectively. RE refers to the relative efficiency of case-cohort estimator, defined by the ratio of mean square errors(MSE) with the full cohort design as a reference.

Table 3: Simulation results for β under model M3 where T_0 follows Exponential distribution with 500 replications.

P_c	β	N	Design	20% cases					10% cases				
				Bias($\hat{\beta}$)	MeanSE	SD($\hat{\beta}$)	95%CP	RE	Bias($\hat{\beta}$)	MeanSE	SD($\hat{\beta}$)	95%CP	RE
0.1	0.1	600	Full cohort	0.001	0.007	0.007	0.950	—	0.002	0.009	0.009	0.956	—
			CC 1:2	0.001	0.008	0.008	0.952	0.790	0.003	0.010	0.010	0.960	0.670
			CC 1:1	0.001	0.009	0.008	0.946	0.679	0.004	0.012	0.011	0.954	0.502
		1200	Full cohort	<0.001	0.005	0.005	0.958	—	0.001	0.006	0.006	0.948	—
			CC 1:2	<0.001	0.005	0.005	0.954	0.800	0.001	0.007	0.007	0.948	0.675
			CC 1:1	<0.001	0.006	0.005	0.956	0.687	0.002	0.008	0.008	0.946	0.561
	0.5	600	Full cohort	0.008	0.056	0.056	0.958	—	0.015	0.060	0.060	0.956	—
			CC 1:2	0.008	0.060	0.060	0.954	0.864	0.020	0.072	0.071	0.948	0.682
			CC 1:1	0.009	0.065	0.065	0.946	0.735	0.022	0.077	0.077	0.954	0.589
		1200	Full cohort	0.009	0.042	0.042	0.948	—	0.009	0.045	0.045	0.946	—
			CC 1:2	0.010	0.047	0.046	0.950	0.812	0.010	0.054	0.055	0.950	0.684
			CC 1:1	0.010	0.049	0.049	0.946	0.722	0.011	0.059	0.060	0.946	0.579
0.4	0.1	600	Full cohort	0.001	0.007	0.007	0.954	—	0.002	0.009	0.009	0.954	—
			CC 1:2	0.001	0.008	0.008	0.956	0.772	0.002	0.011	0.012	0.956	0.614
			CC 1:1	0.001	0.009	0.010	0.956	0.609	0.002	0.013	0.012	0.956	0.490
		1200	Full cohort	<0.001	0.005	0.005	0.954	—	0.001	0.006	0.006	0.950	—
			CC 1:2	<0.001	0.005	0.005	0.954	0.795	0.001	0.008	0.008	0.956	0.626
			CC 1:1	<0.001	0.006	0.005	0.955	0.667	0.001	0.009	0.010	0.958	0.502
	0.5	600	Full cohort	0.008	0.052	0.053	0.974	—	0.018	0.062	0.063	0.963	—
			CC 1:2	0.009	0.056	0.056	0.974	0.860	0.024	0.078	0.079	0.954	0.643
			CC 1:1	0.009	0.061	0.060	0.968	0.728	0.027	0.085	0.085	0.959	0.530
		1200	Full cohort	0.007	0.040	0.041	0.950	—	0.010	0.042	0.043	0.967	—
			CC 1:2	0.008	0.045	0.046	0.948	0.789	0.014	0.050	0.048	0.967	0.666
			CC 1:1	0.008	0.048	0.048	0.946	0.696	0.018	0.056	0.054	0.961	0.534

$\lambda_0 = 0.002, l_0 = 0.5, l_1 = 10, \eta = 1, \sigma_0^2 = 0.25, \sigma_1^2 = 0.25,\ \sigma^2 = 1$. P_c refers to proportion of longitudinal measurements censored due to detection limit. CC 1:2 and CC 1:1 refer to case-cohort study with case-to-control ratio as 1:2 and 1:1, respectively. RE refers to the relative efficiency of case-cohort estimator, defined by the ratio of mean square errors(MSE) with the full cohort design as a reference.

2.5 APPLICATION

We illustrate the proposed estimation procedure with the cytokine data of the Genetic and Inflammatory Markers of Sepsis (GenIMS) study. GenIMS is a multi-center cohort study that enrolled 2320 subjects from the emergency department (ED) with community acquired pneumonia (CAP) in 28 US academic and community hospitals between 2001 and 2003 (Kellum et al. 2007 [25]). One of the primary goals is to identify important inflammatory markers that indicate the risk of severe sepsis and subsequent adverse outcomes, e.g. 90-day mortality. Cytokines including tumor necrosis factors, interleukin-6 (IL6) and interleukin-10 (IL10), were measured for patients admitted to the hospital daily during the first week, and weekly thereafter while they were still in the hospital. We consider IL6 as an example to demonstrate the proposed method, and the objective of our analysis is to investigate whether IL6 changed over time in the first week of hospitalization and how IL6 trajectory was associated with the 90-day mortality.

Among 1895 patients who were confirmed CAP cases admitted to hospitals, 1707 patients had at least one IL6 measurement. All measurements were transformed into a natural log scale to normalize the distribution. The IL6 has two lower detection limits at 2 pg/ml and 5 pg/ml. Censoring proportions from day 1 to day 7 were 14.6%, 21.6%, 29.3%, 32.6%, 34.1%, 35.4%, and 34.4%, respectively. For our analysis, we defined cases as subjects died within 90 days from enrollment. Figure 1 presents the boxplots of log transformed IL6 levels in the first week of hospitalization by 90-day survival status. Overall IL6 concentration was decreased over time for both survivors and nonsurvivors, and nonsurvivors seemed to have higher IL6 at earlier days. The availability of the full cohort data on IL6 allows us to compare the proposed case-cohort estimators to the full cohort estimators. There are about 10% cases among the full cohort. We select subcohorts by different sampling proportions such that the resulting case-cohort studies have 1:1 and 1:2 case-to-control ratios, respectively.

In order to identify an appropriate model for GenIMS data, we consider several joint models for the longitudinal IL6 data and survival data: 2 linear mixed effects models for the natural log transformed IL6, one with random intercept only and the other with both random intercept and slope, 4 AFT models adjusted for baseline characteristics (gender, race

and ICU use) with baseline survival time T_0 following exponential, Weibull, loglogistic, and lognormal distributions, respectively. We fit linear mixed effects models and survival models seperately to abtain the initial values of parameter estimates. We compare model fitting under a joint modelling framework using Akaike information criterions (AICs) and Bayesian information criterions (BICs). The linear mixed effects model with both random intercept and slope fail to provide positive-definite Hessian matrixes in the joint analysis, thus this model is dropped from consideration. The AFT model adjusted for gender, race, and ICU use and with baseline survival time following a Weibull distribution shows superiority in both AIC and BIC (Table 4). Therefore our final joint model is

$$
\begin{aligned}
log(IL6)_{ij} &= l_0 + l_1 t_{ij} + b_{0i} + e_{ij}, && (2.13)\\
T_0 &= \int_0^{T_i} \exp\{\beta(l_0 + l_1 s + b_{0i}) + \eta_r race_i + \eta_g gender_i + \eta_{icu} icu_i\} ds, && (2.14)
\end{aligned}
$$

where $t_i = (t_{i1}, t_{i2}, ..., t_{i7}) = (1, 2, ..., 7)$ and T_0 follows a Weibull distribution.

The results of the analysis, presented in Table 5, are consistent with what we observe from the simulation studies, namely (1) the point estimates under case-cohort designs are close to the full cohort estimates; and (2) the full cohort analysis and case-cohort analysis yield almost identical standard errors for the parameter estimates of the AFT model, while standard errors for the parameter estimates of the mixed effects model are much larger in the case-cohort analysis, especially when the case-cohort sample size is small. In both full cohort and case-cohort analyses, $\hat{\beta}$'s are all positive and $\hat{l}_1$'s are negative, which indicates that the survival time significantly decreases with increasing IL6; and IL6 level slightly decreases over time during the first week of hospitalization. These results are consistent with what we observed in Figure 1.

Table 4: Akaike information criterion (AIC) and Bayesian information criterion (BIC) values for joint analysis of a linear mixed effects model with random intercept and an AFT model adjusted for gender, race and ICU use.

Distribution of baseline survival time (T_0)	AIC	BIC
Exponential	27560	27600
Weibull	27289	27335
Loglogistic	27350	27395
Lognormal	27460	27482

Table 5: Joint analysis of log transformed cytokine IL6 trajectory vs 90-day mortality adjusting for gender, race and ICU use in GenIMS study

	Full cohort			Case-cohort (1:2 ratio)			Case-cohort (1:1 ratio)		
Par	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**	**Estimate**	**SE**	**P value**
				Accelerated Failure Time Model					
β	0.091	0.004	<0.001	0.105	0.005	<0.001	0.112	0.006	<0.001
η_{gender}	-0.003	0.045	0.951	-0.005	0.046	0.917	-0.005	0.046	0.908
η_{race}	-0.014	0.071	0.838	-0.019	0.071	0.794	-0.022	0.072	0.763
η_{icu}	0.053	0.048	0.269	0.059	0.048	0.219	0.063	0.049	0.195
				Mixed effects Model					
l_0	3.685	0.041	<0.001	3.775	0.076	<0.001	3.854	0.096	<0.001
l_1	-0.364	0.008	<0.001	-0.309	0.013	<0.001	-0.290	0.015	<0.001
σ_0^2	1.460	0.029	<0.001	1.531	0.053	<0.001	1.632	0.070	<0.001
σ^2	1.072	0.011	<0.001	1.095	0.020	<0.001	1.128	0.024	<0.001

gender= 1 if male, 0 otherwise; race= 1 if white, 0 otherwise; icu= 1 if yes, 0 otherwise.

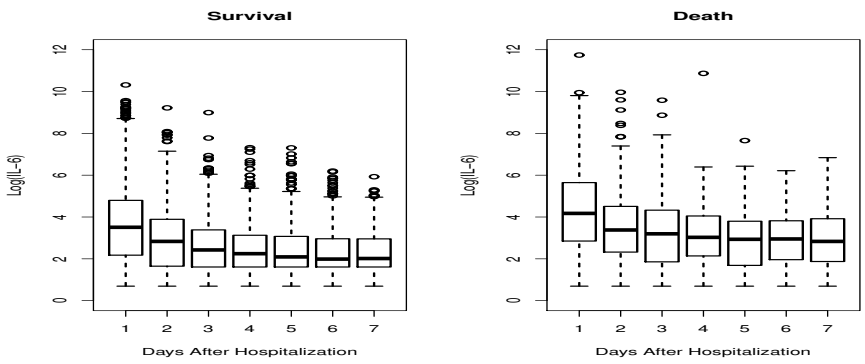

Figure 1: Boxplots for log transformed IL-6 by survival and mortality groups

2.6 DISCUSSION

The case-cohort design is a cost-effective approach to collect data for large cohort studies when exposure variables such as biomarkers or genetic markers are expensive to measure for every subject. We have proposed a modified likelihood-based method for case-cohort designs when accelerated failure time models are used to assess the effect of longitudinal covariate that are subject to measurement error and detection limits. All the event data and easily measured covariates are taken into account in our estimation procedure while classical case-cohort analysis only uses information from the case-cohort sample. We considered a simple case-cohort design where the subcohort is selected by a simple random sampling. A stratified case-cohort design is usually more efficient when the stratification variables are correlated with the outcome of interest. It is straightforward to account for the effect of stratification variables in the AFT model by using stratum specific baseline hazard functions.

When the baseline hazard function in the accelerated failure time model is left unspecified, it becomes challenging to obtain the MLEs. Unlike the Cox model, the baseline hazard function of an AFT model involves both random effects and unknown parameters, which prohibits the use of both the direct maximum likelihood estimation and the nonparametric maximum likelihood estimation with a discrete mass function. Tseng, et. al. (2005) suggested use of a piece-wise constant step function to approximate the baseline hazard

function and presented a Monte Carlo expectation-maximization (EM) algorithm (Dempster, Laird, and Rubin, 1977) to obtain the estimates. Although their estimation procedure performed well as indicated in the simulation studies, it is highly intensive in programming and computation.

With a fully specified baseline hazard function, our method can be easily implemented in SAS Proc NLMIXED procedure and is applicable to nested case-control designs. As demonstrated in the simulation study, the efficiency loss of the proposed case-cohort estimators are relatively small for parameters in the AFT model comparing to full-cohort estimators. In addition, our estimators are almost as efficient as the full-cohort estimators for the covariates that are observed for each individual in the full cohort. However, the efficiency loss can be large for these parameters in the longitudinal model, because information of an survival outcome and baseline covariates are available for every subject in the cohort while longitudinal information is only available for subjects in the case-cohort sample. When a battery of markers is selected from different biological pathways for studying a complex disease, a case-cohort design can be used as an efficient method for initial identification of markers associated with the clinical outcomes while preserving blood samples and reducing the cost.

Interesting extensions of this work include the development of a similar joint-model framework to capture the dependencies between multivariate longitudinal data and time-to-event data, the use of generalized estimation equation method in the estimation procedure, and the control for misclassification error when the longitudinal data is categorical.

3.0 INFERENCE FOR MEAN QUALITY ADJUSTED LIFETIME WITH QUALITY OF LIFE MEASURED REPEATEDLY WITH ERROR

3.1 INTRODUCTION

Survival time is one of the most definitive endpoints used in clinical trials to evaluate the effectiveness of alternative treatments. It is powerful in revealing biological differences in treatment effects, but may be inadequate for deciding the actual benefits for patients. Other end points, such as progression-free survival, physical and psychological health status, and duration of response, are also commonly used for making decisions about medical treatments and the settings of health care programs. Separate analyses of these end points are valid for comparing different aspects of treatments. The choice is straightforward when one treatment shows superiority in all relevant end points. However, the conclusion is difficult to draw when different end points favor different treatments. For instance, a therapy with longer survival time may cause more pain and loss of physical capacity. In this case, the medical decision should be made after evaluating tradeoffs between health benefits and undesirable side effects, in other words, quality of life and length of life. Gelber, Gelman and Goldhirsch (1989) [14] addressed this issue by introducing a quality-of-life-oriented endpoint, time without symptoms of disease and toxicity (TWiST). TWiST is defined by subtracting time with toxic effect of therapy and time with unpleasant symptoms of disease from the overall survival time. It presents the good quality time enjoyed by patients while on treatment.

The concept of quality-adjusted lifetime (QAL) introduced by Goldhirsch et al. (1989) [18] has been considered as an objective measurement that summarizes the quantitative and qualitative health aspects in a unitary and meaningful way. Since its introduction QAL has occupied a significant area of clinical and biomedical research in treatment comparisons,

especially for chronic diseases such as cancer,cardiovascular diseases, and AIDS. The idea is to divide each patient's health experience into several intermediate states, with death at one extreme and perfect health at the other extreme. The time spent in each state is weighted by a utility coefficient ranging from zero to one. The value of the coefficient is decided according to subjective judgments of the quality of life (QOL) that state renders. The perfect healthy state has value one while the absorbing state equivalent to death has value zero. This leads to a utility function over time which takes the value of the utility coefficient of the state occupied at that time. QAL is defined as an integration of this utility function over the survival duration.

In almost all previous works, the health states were defined to be discrete and the number of states were taken to be finite. For discrete, finite states, QAL can be calculated as a sum of time spent at each health state multiplied by the corresponding weight and considered as equivalent to time with perfect health. Gelber, Goldhirsch, and Cavalli (1991) [15] defined QAL by considering three health states: time without symptoms and toxicity (TWiST), time with toxicity (TOX), and survival time after relapse (REL). TWiST is weighted as one and naturally the other two are weighted less. Korn (1993) [29] considered a biased estimator for the distribution of an interpolated area under QOL curve accounting for informative censoring. Zhao and Tsiatis (1997) [62] proposed an estimator of the QAL distribution along with a consistent estimator of its asymptotic variance when the survival time is subject to right censoring. van der Laan and Hubbard (1999) [53] extended this work to incorporate additional time-dependent and time-independent covariates. They showed that their estimator is locally efficient even with dependent censoring. Pradhan and Dewanji (2009) [39] derived a parametric estimator of the QAL distribution for the progressive illness-death models.

In addition to literature developing methods for estimating the distribution of QAL, the mean of QAL has also been studied by many researchers and considered as an important parameter of interest in treatment comparisons. Cole et al. (1993) [8] considered a Cox type parametric regression model to estimate mean QAL using the bootstrap method to obtain the variance estimate. Hwang, Tsauo and Wang (1996) [24] suggested an estimator for the expected quality adjusted survival with quality of life information collected from a cross-sectional survey. Huang and Louis (1999) [23] presented a nonparametric estimation of

the expected quality-adjusted survival time for a general multistate process with incomplete follow-up data. Zhao and Tsiatis (2000) [63] derived a class of consistent and asymptotically normal estimators of mean QAL with right censored data in a more general setting. Wang and Zhao (2007) [57] developed a regression method to investigate the covariate effects on the mean QAL with right-censored data. Andrei and Murray (2007) [2] proposed regression models for the mean of the quality-of-life-adjused restricted survival time using pseudo-observations. Zhao and Wang (2008) [64] considered the empirical likelihood method for the regression model of mean QAL with right-censored data.

In all previous studies the quality function is assumed to be a step function, taking constant value at various states. In most applications, QOL is measured using continuous scales, with scores ranging over some specified intervals. Over the past two decades, several QOL scales have been developed and widely used to describe the change of patients' health status continuously for different diseases. For example, the European Organization for Research and Treatment of Cancer (EORTC) provided QOL scales for use in international clinical trials in oncology [1]. A QOL scale for gastroesophageal reflux disease (GERD) was developed and found to be a valid, reliable, practical, and responsive measure of QOL for patients with GERD by Velanovich et al. (1996) [54]. A reasonable approach to estimate QAL would then be to utilize their QOL measurements in the definition of QAL. Usually the QOL scores are collected repeatedly with error during the course of treatment or the study period. Thus a linear mixed effects model can be applied to describle the trajectory of quality of life scores, which represents the continuous health experience of each patient. We define quality-adjusted lifetime as the survival time adjusted for the continually changing health status through a link function that maps the quality-of-life scores to the interval [0,1]. To improve the estimation of parameters in the linear mixed effects model, we apply joint modelling method to longitudinal quality-of-life scores and the survival time.

It is well known that the Kaplan-Meier estimator for the survival distribution of QAL is biased due to the induced informative censoring (Gelber, et al., 1989 [14]). To eliminate this bias, Glasziou et al. (1990) [17] propose a consistent estimator for mean QAL for progressive state models, in which patients' health state changes from one to another following a fixed sequence. Their estimator of mean QAL is calculated as the weighted sum of areas under the

separate Kaplan-Meier survival curves for each health state, but no consistent estimator for the variance is provided. Furthermore, right censoring is very common in clinical trials due to the lost to follow-ups and the restriction on the length of studies. Therefore, the complete health experience, as well as the survival time, cannot be observed for all patients. With censored data, a consistent nonparametric estimate of mean QAL over the entire health history cannot be obtained. Zhao and Tsiatis (2000) [63] suggest calculating mean QAL restricted to a certain length of time, which is usually determined by the length of study, and applying the general representation theorem for the missing data process developed by Robins and Rotnitzky (1992) [41] and Robins, et al. (1994) [42] to derive a class of consistent and asymptotically normal estimators for mean QAL and their asymptotic variance in a general setting. We borrow the basic idea from them and derive a consistent and asymptotically normal estimator for mean QAL and its asymptotic variance where the utility function is esimated from the data based on the QOL scores.

In Section 3.2 and 3.3 we introduce the notation and estimation of mean QAL with its asymptotic variance. Section 3.4 describles the simulation studies we conducted to evaluate our estimators. Section 3.5 provides some discussions and possible direction of future works.

3.2 METHOD

3.2.1 Set-up

To begin, we introduce some notation and the data structure. For the ith individual, let T_i^* be the potential event time, such as time to death, time to disease progression or time to infection, and C_i be the potential right censoring time, such as time to lost to follow-up or time to the end of the study, respectively. Due to the presence of censoring, we have to consider quality adjusted lifetime within a restricted length of time. Let L denote a time after which a reasonable proportion of individuals are still being followed. L is usually determined by the length of study. Therefore, the survival time would be defined as $T_i = min(T_i^*, L)$. We define QAL for the ith patient as

$$U_i = \int_0^{T_i} Q(W_i(t))dt, \tag{3.1}$$

where Q is a monotone link function that maps the QOL process $W(.)$ to the interval $[0, 1]$, where 0 indicates the lowest quality of life such as death and 1 indicates the best quality, e.g. perfect health. The function Q is usually known and parameterized, for example, using an inverse logit function. Thus $Q(W_i(t))$ for an inverse logit link can be written as

$$Q(W_i(t)) = \begin{cases} 0 & \text{if } t > T_i \\ logit^{-1}\{qW_i(t)\} & \text{if } t \leq T_i \end{cases}, \tag{3.2}$$

where the scale parameter q changes the scale of the QOL scores in the QAL function.

3.2.2 Observed Data

Due to right censoring survival times are not always observed completely. Denote the observed event time and event indicator as Y_i and $\triangle_i$, respectively, where $Y_i = min(T_i, C_i)$ and $\triangle_i = 1$, if $T_i \leq C_i$, and 0 otherwise. QOL scores are measured usually at fixed time points until the terminal event. Assume the quality of life scores $W_i = \{W_{i1}, W_{i2}, \ldots, W_{im_i}\}$ are obtained from the ith individual at time $t_i = \{t_{i1}, t_{i2}, \ldots, t_{im_i}\}$, for $t_{im_i} \leq T_i$. Let $X_i(t)$ and $Z_i(t)$ be possible time-dependent covariate vector process that are potentially associated with mean and subject-specific QOL score and V_i be a baseline covariate vector. Then the observed data for the ith individual is $\{Y_i, \triangle_i, X_i, Z_i, V_i, W_i, t_i\}$, where $X_i = \{X_{i1}, X_{i2}, \ldots, X_{im_i}\}$ and $Z_i = \{Z_{i1}, Z_{i2}, \ldots, Z_{im_i}\}$ are the realizations of $X_i(t)$ and $Z_i(t)$, respectively, at time $t_i = \{t_{i1}, t_{i2}, \ldots, t_{im_i}\}$. Our goal is to draw inferences about the mean QAL, $\mu = E(U) = E\left[\int_0^T Q(W(t))dt\right]$.

3.3 INFERENCE

3.3.1 Models

As mentioned in Section 3.2.1, we define the QAL for the ith patient as

$$U_i = \int_0^{T_i} Q(W_i(t))dt.$$

However, the QOL process $W_i(t)$ is unknown for all t. It is measured with error only at fixed timepoints as the QOL score W_{ij} for the ith individual at time t_{ij}, for $t_{ij} \leq Y_i$. Given the dependence between the quality of life and the occurence of a terminal event, we propose to estimate the QOL process $W_i(t)$ by jointly modeling the observed QOL scores and the survival data. More specifically, we can define marginal models for the QOL scores and the survival data with shared random effects and estimate the parameters of interest by maximizing the joint likelihood of these two marginal models. We propose to use the marginal models specified in Sections 3.3.1.1 and 3.3.1.2.

3.3.1.1 Model for the longitudinal QOL score We propose using a linear mixed effects model for the observed QOL score W_{ij}. The proposed model can be writen as

$$\begin{aligned} W_{ij} &= W_i(t_{ij}) + G_i(t_{ij}) + e_{ij} \\ &= f(X_i(t_{ij}); \beta) + g(Z_i(t_{ij}); b_i) + e_{ij}, \end{aligned} \tag{3.3}$$

where $W_i(t) = f(X_i(t); \beta)$ describes the trajactory of mean QOL scores modeled as a function of the fixed effect β and its design vector $X_i(t)$ that might contain individual baseline covariates and the measurement times t_{ij}; $G_i(t) = g(Z_i(t); b_i)$ represents the within subject variation modeled as a function of the random effect b_i and its subject specific design vector $Z_i(t)$ that also might be time-dependent; and e_{ij} is the measurement error, assumed to be independently normally distributed with mean zero and variance σ^2. We assume the random effect b_i is normally distributed with mean zero and covariance matrix Σ and it is independent of the measurement error e_{ij}. This model not only allows the assessment of change in the QOL process over time using fixed effects, but also accounts for within subject correlation through subject-specific random effects.

3.3.1.2 Model for the survival time We propose using an accelerated failure time (AFT) model to describe the survival data due to its appealing interpretation and robustness to non-proportionality of hazards. The QOL score is closely related to the survival time because it is defined to be a summary statistic of both the quantitative and qualitative health aspects. We incorparate the mean QOL score as an importent covariate in the survival model and the propose AFT model can be writen as

$$\begin{aligned} T_0 &= \int_0^{T_i} \exp\{\alpha W_i(s) + \eta V_i\} ds \\ &= \int_0^{T_i} \exp\{\alpha f(X_i(s);\beta) + \eta V_i\} ds, \end{aligned} \tag{3.4}$$

where T_0 is referred to as the baseline survival time since it is the survival time for subjects with all covariates equal to zero; $W_i(s) = f(X_i(s);\beta)$ is a function defined the same as in Section 3.3.1.1, where $X_i(s)$ is a possible time-dependent design vector for fixed effect β; and V_i is a vector of baseline covariates that affect the survival time. And the hazard function can be expressed by

$$\lambda(t|\bar{f}(X_i(t);\beta), V_i) = \lambda_0 \left(\exp\{\eta V_i\} \int_0^t \exp\{\alpha f(X_i(s);\beta)\} ds \right) \exp\{\alpha f(X_i(t);\beta) + \eta V_i\}, \tag{3.5}$$

where $\bar{f}(X_i(t);\beta) = \{f(X_i(s);\beta), s \le t\}$ is the mean QOL score trajectory up to time t, and λ_0 is the hazard function of T_0 usually referred to as baseline hazard function.

3.3.1.3 Estimation of β The parameterization of the QOL process $W_i(t)$ in terms of fixed effect β allows us to estimate it. We obtain the maximum likelihood estimator $\hat{\beta}$ of β by jointly fitting Models (3.3) and (3.4) and maximizing the joint likelihood. Then, if T_i's were known, the QAL can be estimated as

$$U_i(\hat{\beta}) = \int_0^{T_i} Q(W_i(t)) dt = \begin{cases} 0 & \text{if } t > T_i \\ \int_0^{T_i} logit^{-1}\{q f(X_i(t_{ij});\hat{\beta})\} dt & \text{if } t \le T_i \end{cases}, \tag{3.6}$$

and the corresponding mean QAL could be estimated as

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} U_i(\hat{\beta}). \tag{3.7}$$

However, T_i's are not all known, and hence the above estimator is not of practical use. One, therefore, may choose to predict T_i based on Model (3.4) and replace the predicted T_i in the above equation to obtain a predicted U_i, namely $\hat{U}_i(\hat{\beta})$ and then average over all subjects to construct an estimator of μ. We refer to this estimator as

$$\hat{\mu}^P = \frac{1}{n}\sum_{i=1}^{n}\hat{U}_i(\hat{\beta}). \tag{3.8}$$

Prediction of survival times are often cumbersome, specially, in the presence of time-dependent covariates. Therefore, we will use inverse-probability weighting method to account for the censored survival times in Equation (3.6). We describe the detailed procedure and the corresponding large sample properties in the next section.

3.3.2 Estimation of mean QAL

Inverse probability weighting has been used extensively in the analysis of survival data. For instance, Rotnitzky and Robins (2005) [43] discussed inverse probability weighted augmented estimation in survival studies, which provides a robust approach to model misspecification; and Xie and Liu (2005) [59] proposed adjusted Kaplan-Meier estimator with inverse probability of treatment weighting for survival data. Zhao and Tsiatis (2000) [63] used this in the QAL setting to derive the estimator for mean QAL. Following their work, we propose to use a simple weighted estimator for mean QAL:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\frac{\triangle_i\hat{U}_i(\hat{\beta})}{\hat{K}(Y_i)}, \tag{3.9}$$

where $\hat{U}_i(\hat{\beta}) = \int_0^{Y_i} Q(W_i(t))dt$ is defined the same way as $U_i(\hat{\beta})$ in Equation (3.6) except that T_i is replaced by the observed Y_i, and $\hat{K}(.)$ is the Kaplan-Meier estimator for the survival distribution function of the censoring time C_i. With data $\{Y_i = min(T_i, C_i), \triangle_i, i = 1, 2, \ldots, n\}$, $\hat{K}(t)$ can be estimated as

$$\hat{K}(t) = \prod_{u\leq t}\left\{1 - \frac{dN^c(u)}{Y(u)}\right\}, \tag{3.10}$$

where $N^c(u) = \sum_i I(Y_i \leq u, \triangle_i = 0)$, and $Y(u) = \sum I(Y_i \geq u)$. Fleming and Harrington (1991) [13] showed that $\hat{K}(t)$ is a uniformly consistent estimator of $K(t)$ over $[0, L]$. Following

the procedure described in Section 3.3.1.3, $\hat{\beta}$ is a consistent estimator for β, the simple weighted estimator $\hat{\mu}$ in Equation (3.9) for mean QAL can be shown to be a consistent estimator, as long as that (i) given W_{ij}, the censoring is independent of survival; (ii) the mean model of the QOL score, $E(W_{ij}) = W_i(t_{ij}) = f(X_i(t_{ij}); \beta)$, and the AFT model described in Section 3.3.1.2 are correctly specified. We first note that when $\triangle_i = 1$, $\hat{U}_i(\hat{\beta}) = U_i(\hat{\beta})$ and $\hat{K}(Y_i) = \hat{K}(T_i)$, therefore, $\hat{\mu}$ in Equation (3.9) can be written equivalently as

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \frac{\triangle_i U_i(\hat{\beta})}{\hat{K}(T_i)}. \tag{3.11}$$

To motivate the construction of this estimator we first show that when $\hat{K}(.)$ and $\hat{\beta}$ are known, i.e. $\hat{K}(T_i) = K(T_i)$, $\hat{\beta} = \beta_0$. The estimator $\hat{\mu}^0 = \frac{1}{n}\sum_{i=1}^{n} \frac{\triangle_i U_i(\beta_0)}{K(T_i)}$ is unbiased for μ. Specifically,

$$\begin{aligned}
E\{\hat{\mu^0}\} &= E\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{\triangle_i U_i(\beta_0)}{K(T_i)}\right\} = \frac{1}{n}\sum_{i=1}^{n} E\left[E\left\{\frac{\triangle_i U_i(\beta_0)}{K(T_i)}|T_i, W_i\right\}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} E\left[\frac{U_i(\beta_0)}{K(T_i)} E\{I(C_i \geq T_i)|T_i, W_i\}\right] \\
&= \frac{1}{n}\sum_{i=1}^{n} E\left[\frac{U_i(\beta_0)}{K(T_i)} K(T_i)\right] \\
&= E\left[\frac{1}{n}\sum_{i=1}^{n} U_i(\beta_0)\right] = E[U(\beta_0)] = E[U] = \mu.
\end{aligned} \tag{3.12}$$

One then would expect that the proposed estimator $\hat{\mu}$ defined in Equation (3.9), will be a consistent estimator of μ, since $\hat{\mu}$ only replaces the unknown β and K in $\hat{\mu}^0$ by their uniformly consistent estimators.

3.3.3 Consistency of $\hat{\mu}$

Since $\hat{\beta}$ in obtained through the optimazation of the joint likelihood, $\hat{\beta}$ is a consistent and asymptotically normal estimator. Let

$$L(\beta) = \frac{1}{n}\sum_{i=1}^{n} l_i(\beta|Y_i, X_i, W_i, V_i)$$

where $l_i(\beta|Y_i, X_i, W_i, V_i)$ is the log joint likelihood function of Models (3.3) and (3.4). We denote the first derivative of $L(\beta)$ as $L'(\beta)$ and the second derivative as $L''(\beta)$. Then by

the definition of maximum likelihood estimate, $L'(\hat{\beta}) = 0$. By the Mean Value Theorem, we know

$$0 = L'(\hat{\beta}) = L'(\beta_0) + (\hat{\beta} - \beta_0)^T L''(\hat{\beta}_1),$$

for some $\hat{\beta}_1 \in [\hat{\beta}, \beta_0]$ and $L''(\hat{\beta}_1) \approx E[L''(\beta_0)]$. Therefore, we can write

$$\hat{\beta} - \beta_0 = \frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i) + o_p(1)$$

where ψ_i is proportional to the score equation of β_0, specifically

$$\psi_i(\beta_0 | Y_i, X_i, W_i, V_i) = -E[l''(\hat{\beta}_1 | Y_i, X_i, W_i, V_i)]^{-1} \{l'(\beta_0 | Y_i, X_i, W_i, V_i)\}^T. \tag{3.13}$$

Now $U_i(\hat{\beta})$ can be approximated using a Taylor's series expansion as follow,

$$\begin{aligned} U_i(\hat{\beta}) &\approx U_i(\beta_0) + (\hat{\beta} - \beta_0)^T U_i'(\beta_0) \\ &= U_i(\beta_0) + \left\{ \frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i) + o_p(n^{-\frac{1}{2}}) \right\}^T U_i'(\beta_0), \end{aligned}$$

where $U_i'(\beta_0)$ is the first derivative vector of $U_i(\beta)$ with respect to β evaluated at β_0. Now,

$$\begin{aligned} \hat{\mu} - \mu &= \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i(\hat{\beta})}{\hat{K}(T_i)} \\ &\approx \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i}{\hat{K}(T_i)} \left\{ U_i(\beta_0) + \left\{ \frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i) + o_p(n^{-\frac{1}{2}}) \right\}^T U_i'(\beta_0) \right\} - \mu \\ &= \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i(\beta_0)}{\hat{K}(T_i)} - \mu \\ &+ \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i'(\beta_0)}{\hat{K}(T_i)} \right\}^T \frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i) \\ &+ \left\{ o_p(n^{-\frac{1}{2}}) \right\}^T \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i'(\beta_0)}{\hat{K}(T_i)}. \end{aligned} \tag{3.14}$$

The first term in Equation (3.14) is equal to

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i \{U_i(\beta_0) - \mu\}}{K(T_i)} - \frac{1}{n} \sum_{i=1}^{n} \triangle_i \{U_i(\beta_0) - \mu\} \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i) K(T_i)}. \tag{3.15}$$

The first term is a weighted average of n independent, identically distributed random variables, and its expected value is zero as showed in Equation (3.12). So the first term in

Equation (3.15) converges to zero in probability. By assumption, $T_i \leq L$ and $K(L) > 0$, the second term is bounded from above by

$$\frac{sup_{u \leq L} |\hat{K}(u) - K(u)|}{\hat{K}(L) - K(L)}. \tag{3.16}$$

Since $\hat{K}(u)$ converges uniformly to $K(u)$ for $u \leq L$ as showed by Fleming and Harrington (Ch. 6) [13], Equation (3.16) converges to zero in probability. Therefore, the second term converges to zero in probability. Thus Equation (3.15) converges to zero in probability. The second term in Equation (3.14) is equal to

$$\begin{aligned} &\left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i'(\beta_0)}{K(T_i)} \right\}^T \frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i) \\ &+ \left\{ \frac{1}{n} \sum_{i=1}^{n} \triangle_i U_i'(\beta_0) \frac{\hat{K}(T_i) - K(T_i)}{\hat{K}(T_i) K(T_i)} \right\}^T \frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i), \end{aligned} \tag{3.17}$$

where $\frac{1}{n} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i)$ converges to zero in probability since ψ_i is proportional to the score function, $\frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i'(\beta_0)}{K(T_i)}$ converges to a constant vector, C, the expected value of $U'(\beta_0)$, and the second term in Equation (3.17) converges to zero in probability since it is also bounded from above by Equation (3.16). Therefore Equation (3.17) converges to zero in probability. Following the same logic, the third term in Equation (3.14) also converges to zero in probability since $o_p(n^{-\frac{1}{2}}) = n^{-\frac{1}{2}} o_p(1)$ converges to zero. This shows Equation (3.14) converges to zero in probability. So the proposed estimator is a consistent estimator for the mean QAL.

Furthermore, we investigate the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\mu} - \mu)$. According to Equation (3.14), we can write

$$\begin{aligned} n^{\frac{1}{2}}(\hat{\mu} - \mu) &= n^{-\frac{1}{2}} \sum_{i=1}^{n} \left\{ \frac{\triangle_i U_i(\beta_0)}{\hat{K}(T_i)} - \mu \right\} \\ &+ \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{\triangle_i U_i'(\beta_0)}{\hat{K}(T_i)} \right\}^T n^{-\frac{1}{2}} \sum_{i=1}^{n} \psi_i(\beta_0 | Y_i, X_i, W_i, V_i) \\ &+ \{o_p(1)\}^T n^{-1} \sum_{i=1}^{n} \frac{\triangle_i U_i'(\beta_0)}{\hat{K}(T_i)}. \end{aligned} \tag{3.18}$$

In order to find the asymptotic distribution of $n^{\frac{1}{2}}(\hat{\mu}-\mu)$, we need to find the asymptotic distribution of the above three terms. Let $H_1(\beta_0) = n^{-\frac{1}{2}}\left\{\sum_{i=1}^{n}\frac{\triangle_i U_i(\beta_0)}{\hat{K}(T_i)} - n\mu\right\}$, $H_2(\beta_0) = \frac{1}{n}\sum_{i=1}^{n}\frac{\triangle_i U_i'(\beta_0)}{\hat{K}(T_i)}$ and $H_3(\beta_0) = n^{\frac{1}{2}}\sum_{i=1}^{n}\psi_i(\beta_0|Y_i, X_i, W_i, V_i)$. Equation (3.18) can be written as

$$n^{\frac{1}{2}}(\hat{\mu}-\mu) = H_1(\beta_0) + \{H_2(\beta_0)\}^T H_3(\beta_0) + o_p(1)^T H_2(\beta_0). \tag{3.19}$$

We can rewrite $H_1(\beta_0)$ as

$$H_1(\beta_0) = n^{-\frac{1}{2}}\left[\sum_{i=1}^{n}\frac{\triangle_i U_i(\beta_0)}{K(T_i)} - \sum_{i=1}^{n}\triangle_i U_i(\beta_0)\left\{\frac{\hat{K}(T_i)-K(T_i)}{\hat{K}(T_i)K(T_i)}\right\} - n\mu\right]. \tag{3.20}$$

Let $\lambda^c(u)$ be the hazard function for the censoring distribution, and the corresponding martingale process $M_i^c(u)$ can be expressed as

$$M_i^c(u) = N_i^c(u) - \int_0^u \lambda^c(t)Y_i^*(t)dt,$$

where $N_i^c(u) = I(Y_i \le u, \triangle_i = 0)$ and $Y_i^*(u) = I(Y_i \ge u)$. Let $M^c(u) = \sum M_i^c(u)$, $N^c(u) = \sum N_i^c(u)$, and $Y(u) = \sum Y_i(u)$. According to Robins and Rotnitzky (1992) [41] that

$$\frac{\triangle_i}{K(T_i)} = 1 - \int_0^\infty \frac{dM_i^c(u)}{K(u)},$$

Gill (1980) [16] that

$$\frac{\hat{K}(T_i)-K(T_i)}{K(T_i)} = -\int_0^{T_i}\frac{\hat{K}(u)}{K(u)}\frac{dM^c(u)}{Y(u)},$$

and

$$n^{-1}Y(u) = \hat{K}(u)\hat{S}(u),$$

where $\hat{S}(u)$ is the Kaplan-Meier estimate for $S(u) = pr(T > u)$, Equation (3.20) can be written as

$$n^{-\frac{1}{2}}\sum_{i=1}^{n}\{U_i(\beta_0)-\mu\} - n^{-\frac{1}{2}}\sum_{i=1}^{n}\int_0^\infty \frac{dM_i^c(u)}{K(u)}\{U_i(\beta_0) - \hat{G}(U,u)\}, \tag{3.21}$$

where $\hat{G}(U,u) = \frac{1}{n\hat{S}(u)} \sum_{i=1}^{n} \frac{\triangle_i U_i(\beta_0) I(T_i \geq u)}{\hat{K}(T_i)}$. Since $U_i(\beta_0)$ is $F(0)$, the two terms in Equation (3.21) are uncorrelated. Following the derivation in Zhao and Tsiatis (1997) [62], we rewrite Equation (3.21) as

$$n^{-\frac{1}{2}} \sum_{i=1}^{n} \{U_i(\beta_0) - \mu\} - n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^{\infty} \frac{dM_i^c(u)}{K(u)} \{U_i(\beta_0) - G(U,u)\}$$
$$-n^{-\frac{1}{2}} \sum_{i=1}^{n} \int_0^{\infty} \frac{dM_i^c(u)}{K(u)} \{G(U,u) - G^*(U,u)\} - n^{-\frac{1}{2}} \{\hat{\mu}^0 - \mu\} \sum_{i=1}^{n} \int_0^{\infty} \frac{dM_i^c(u)}{K(u)\hat{S}(u)} \quad (3.22)$$

where

$$G(U,u) = \frac{1}{S(u)} E\{U_i(\beta_0) I(Y_i \geq u)\}$$

and

$$G^*(U,u) = \frac{1}{\hat{S}(u)} \left\{ \mu - n^{-1} \sum_{i=1}^{n} \frac{\triangle_i U_i(\beta_0) I(T_i < u)}{\hat{K}(T_i)}. \right\} \quad (3.23)$$

From the central limit theorem. the first term in Equation (3.22) is asymptotically normally distributed with mean zero and variance $Var(U_i(\beta_0) - \mu)$. From the martingale central limit theorem (Fleming and Harrington, 1991 [13]), the second term also converges to a normal distribution with mean zero and variance

$$E \int_0^{\infty} [U_i(\beta_0) - G(U(\beta_0), u)]^2 I(Y_i \geq u) \frac{\lambda^c(u)}{K(u)} du. \quad (3.24)$$

According to Zhao and Tsiatis (1997) [62], the third term converges to a constant, zero, in probability. The last term in Equation (3.22) also converges to zero since $\hat{\mu}^0 - \mu$ converges to zero in probability as shown in Equation (3.12), and

$$\sum_{i=1}^{n} \int_0^{\infty} \frac{dM_i^c(u)}{K(u)\hat{S}(u)}$$

converges to a normal distribution by the martingale central limit theorem. Then by Slutsky's theorem, $H_1(\beta_0)$ is asymptotically normally distributed with mean zero and variance

$$\sigma_{H_1} = Var(U_i(\beta_0) - \mu) + E \int_0^{\infty} [U_i(\beta_0) - G(U(\beta_0), u)]^2 I(Y_i \geq u) \frac{\lambda^c(u)}{K(u)} du. \quad (3.25)$$

This variance can be estimated

$$\hat{\sigma^2_{H_1}} = n^{-1} \sum_{i=1}^{n} \frac{\triangle_i \{U_i(\hat{\beta}) - \hat{\mu}\}^2}{\hat{K}(Y_i)} + n^{-1} \int_0^{\infty} \frac{dN^c(u)}{\hat{K}(u)^2} \{\hat{G}(U(\hat{\beta})^2, u) - \hat{G}(U(\hat{\beta}), u)^2\}. \quad (3.26)$$

Following the same argument, $H_2(\beta_0)$ can be written as

$$\begin{aligned} H_2(\beta_0) &= \frac{1}{n}\sum_{i=1}^{n}\frac{\triangle_i U_i'(\beta_0)}{K(T_i)} - \frac{1}{n}\sum_{i=1}^{n}\triangle_i U_i'(\beta_0)\frac{\hat{K}(T_i)-K(T_i)}{\hat{K}(T_i)K(T_i)} \\ &= n^{-1}\sum_{i=1}^{n}\left[\frac{dU_i(\beta_0)}{d\beta_0} - \int_0^{\infty}\frac{dM_i^c(u)}{K(u)}\left\{\frac{dU_i(\beta_0)}{d\beta_0} - \frac{1}{S(u)}E(\frac{dU_i(\beta_0)}{d\beta_0}I(Y_i \geq u))\right\}\right], \end{aligned}$$

which is a sum of independent and identically distributed samples. Thus $H_2(\beta_0)$ converges to a constant vector, C, that

$$C = E\left[\frac{dU_i(\beta_0)}{d\beta_0}\right]. \tag{3.27}$$

From Equation (3.13), we can write $H_3(\beta_0)$ as

$$\begin{aligned} H_3(\beta_0) &= n^{-\frac{1}{2}}\sum_{i=1}^{n}\psi_i(\beta_0|Y_i, X_i, W_i, V_i) \\ &= n^{-\frac{1}{2}}\sum_{i=1}^{n} -E[l''(\beta_0|Y_i, X_i, W_i, V_i)]^{-1}\{l'(\beta_0|Y_i, X_i, W_i, V_i)\}^T. \end{aligned} \tag{3.28}$$

As sample size goes to infinity, Equation (3.28) can be written as

$$\begin{aligned} n^{\frac{1}{2}}\sum_{i=1}^{n} -E[l''(\beta_0|Y_i, X_i, W_i, V_i)]^{-1}\{l'(\beta_0|Y_i, X_i, W_i, V_i)\}^T \\ = n^{\frac{1}{2}} - E[L''(\beta_0)]^{-1}L'(\beta_0)^T, \end{aligned} \tag{3.29}$$

where $L'(\beta_0)$ is the score function for β_0 that converges to a normal distribution with mean zero and variance $I(\beta_0)$, which is the corresponding expected Fisher information matrix, and $-E[L''(\beta_0)]^{-1}$ equals $[I(\beta_0)]^{-1}$. So Equation (3.29) converges to a normal distribution with mean zero and variance

$$\sigma_{H_3} = [I(\beta_0)]^{-1}. \tag{3.30}$$

To sum up, when n goes to infinite, the first term in Equation (3.18) converges to a normal distribution, the second term also converges to a normal distribution since it becomes a product of a constant and a normally distributed variable, and the third term converges to zero since $o_p(n) = no_p(1)$ converges to zero. These three terms are not correlated with each other. Therefore, $n^{\frac{1}{2}}(\hat{\mu} - \mu)$ is asymtotically normally distributed with mean zero and variance

$$\sigma_{H_1} + C\sigma_{H_3}C^T, \tag{3.31}$$

where the form of σ_{H_i}, C, and σ_{H_3} are showed in Equations (3.25), (3.27), and (3.30), respectively. And this variance can be estimated by

$$\hat{\sigma_{H_1}} + H_2(\hat{\beta})\hat{\sigma_{H_3}}H_2(\hat{\beta})^T. \tag{3.32}$$

3.4 SIMULATION STUDIES

We evaluate the performance of our proposed mean quality adjusted lifetime estimator and its corresponding variance estimator in a number of simulation studies. We generate the quality of life score W_{ij} for individual i at time t_{ij} from the following mixed effects model,

$$\begin{aligned} W_{ij} &= \beta_0 + \beta_1 t_{ij} + b_{0i} + b_{1i}t_{ij} + e_{ij}, \\ \mathbf{b}_i &= \begin{pmatrix} b_{0i} \\ \\ b_{1i} \end{pmatrix} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), e_{ij} \sim \mathcal{N}(\mathbf{0}, \sigma^{\mathbf{2}}), \end{aligned} \tag{3.33}$$

where $\beta_0 = 2$, $\beta_1 = -3$, $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \\ 0 & 0.5 \end{pmatrix}$, $\sigma = 1$ and assume 10 measurements for each individual at time $t_i = \{t_{ij}, j = 1, 2, \ldots, 10\} = \{1, 2, \ldots, 10\}$. Considering the dependency between the quality of life score and the event time, the following accelerated failure time model is used to generate the survival data,

$$T_0 = \int_0^{T_i} exp\left\{\alpha(\beta_0 + \beta_1 t) + \eta V\right\} dt, \tag{3.34}$$

where T_0 is the baseline survival time, which is the survival time for an individual with all covariates equal to zero, follows a weibull distribution with a shape parameter $a = 1.5$ and a scale parameter $b = 80$, $\alpha = -0.2$, $f(t;\beta) = \beta_0 + \beta_1 t$ is the function of the population mean of the quality of life score with β_0 and β_1 specified the same as in Model (3.33), and $\eta = 1$. We consider 3 settings for the baseline covariate V in Model (3.34): (i) $V = V_1$, where V_1 follows a binomial distribution with probability $p = 0.5$; (ii) $V = V_2$, where V_2 follows a standard normal distribution; (iii) V includes both V_1 and V_2 with corresponding coefficient η_1 and

η_2, where V_1 and V_2 follow distributions as specified before. For each survival model, we generate uncensored survival data with known β for a very large population (n=10000000) first. The truncation time, L, is chosen to be the third quartile of the survival time of this large population. This helps to make sure that a certain proportion of the population are still followed after L in our simulation studies. Then the true mean of quality adjusted lifetime for a restricted time period $[0, L]$ is calculated as the average of the quality adjusted lifetime defined as

$$U_i = \int_0^{T_i} logit^{-1}\{q(\beta_0 + \beta_1 t)\}dt, \tag{3.35}$$

where $q = 0.001$ is a scale parameter. The censoring time C_i is generated from weibull distributions with shape and scale parameters chosen appropriately such that the resulting censoring proportion is about 20% and 40% . In order to more efficiently make use of all available information, we estimate β and the variance of its estimator by applying the joint modelling technique to Models (3.33) and (3.34). Then the proposed estimator of the mean quality adjusted lifetime is obtained by Equation (3.9) with $\hat{\beta}$ being the point estimator of β, and its asymptotic variance is estimated by Equation (3.32). When β is assumed to be known, the asymptotic variance of our proposed estimator can be calculated by Equation (3.26).

We carry out the simulation studies under 8 scenarioes using survival models with different baseline covariates as specified in Tables 6 and 7. In order to investigate the robustness of our method to model misspecification, we fit survival models with a list of covariates different from the true model and compare the results with that of the correct model. We conducted 500 simulations under each setting with sample sizes N=200, 400, and 800.

The proposed estimators for the mean QAL have minimal biases (< 0.02) under all settings. The bias decreases with increasing sample size. The average of the estimated standard errors are very close to the means of the empirical standard derivations of the mean QAL based on 500 simulation samples. The coverage probabilities of the 95% confidence intervals are at nominal level. The biases of the proposed estimator increase slightly when the survival models are misspecified, but still remain small, especially when the sample size is large. In summary, the proposed estimator for the mean QAL is approximately unbiased. Its asymptotic variance formula yields accurate results.

3.5 CONCLUSION AND DISCUSSION

In this chapter, we propose an estimator for the mean quality adjusted lifetime with a continuous quality of life function. This estimator takes into account the continuity of people's health experience, and thus better summarizes the quantitative and qualitative health aspects of people. With the help of well-developed quality of life scales, we are able to describe people's continuously changing health status by the trajectory of quality of life score defined as our quality of life function. We realize the possible biases that might be introduced by directly using the scores since the measurement is usually taken infrequently and with error during the course of studies. In order to avoid this bias and given the dependency between the quality of life status and the survival time, we consider jointly modeling the quality of life score and the survival time using a mixed effects model and a survival model as an ideal approach. We find the most appropriate values for the parameters in the quality of life function by maximizing the joint likelihood, and then calculate the quality adjusted lifetime by integrating the quality of life function over the observed survival time. We suggest an estimator for the mean quality adjusted lifetime in a restricted time period for right censored data, and derive its asymptotic variance. The simulation studies show that our proposed estimator is unbiased and robust to model misspecification, and our variance formula is correct. Our proposed estimator enables a comprehensive comparison of treatments that may benefit people's health in different ways or be applied to different populations in clinical trials.

However, the consistency of our estimator depends on the assumption the censoring time is independent of the survival time and the health experience. In situations that the censoring depends on the survival time, we may consider the use of imputation techniques for missing data, sensitivity analyses to mimic best and worst-case scenarios and use of the drop-out event as a study end-point [46]. If the censoring depends on the health experience, theoretically we can model the actual distribution of the censoring time and substitute the Kaplan-Meier estimator for the censoring time in the denominator of our proposed estimator with it. This should help to maintain the good performance of the proposed estimator. These studies will be subjects of future research.

Table 6: Simulation results for the estimator of the mean quality adjusted lifetime, $\hat{\mu}$, and its asymptotic standard error for 20% censored data based on 500 replications.

True model covariate distribution	True μ	N	Baseline covariate (Fit)	Bias($\hat{\mu}$)	SD($\hat{\mu}$)	MeanSE	95% CP
$V_1 \sim BIN(0.5)$	8.577	200	Known β	0.003	0.097	0.099	0.950
			Binary	0.003	0.096	0.099	0.954
			Continuous	0.010	0.100	0.099	0.946
		400	Known β	0.007	0.072	0.070	0.922
			Binary	0.007	0.072	0.070	0.926
			Continuous	0.015	0.078	0.079	0.921
		800	Known β	0.001	0.050	0.050	0.936
			Binary	0.001	0.050	0.050	0.942
			Continuous	0.006	0.053	0.054	0.938
$V_2 \sim N(0,1)$	8.980	200	Known β	0.006	0.109	0.109	0.940
			Continuous	0.007	0.108	0.109	0.934
			Binary	0.011	0.111	0.113	0.933
		400	Known β	0.011	0.077	0.078	0.942
			Continuous	0.009	0.087	0.091	0.956
			Binary	0.013	0.091	0.096	0.951
		800	Known β	<0.001	0.053	0.055	0.950
			Continuous	0.001	0.060	0.066	0.972
			Binary	0.009	0.063	0.066	0.961
V_1 and $V2$	8.780	200	Known β	0.003	0.053	0.052	0.947
			Binary and Continuous	0.003	0.054	0.055	0.946
			Binary	0.012	0.057	0.056	0.939
			Continuous	0.007	0.056	0.056	0.941
		400	Known β	0.003	0.038	0.037	0.948
			Binary and Continuous	0.003	0.038	0.040	0.955
			Binary	0.014	0.042	0.043	0.952
			Continuous	0.005	0.039	0.041	0.951
		800	Known β	0.001	0.027	0.026	0.943
			Binary and Continuous	0.001	0.028	0.031	0.973
			Binary	0.009	0.031	0.033	0.955
			Continuous	0.004	0.032	0.034	0.962

$\beta_0 = 2$, $\beta_1 = -3$, $\sigma_{11} = 1$, $\sigma_{12} = \sigma_{21} = 0$, $\sigma_{22} = 0.5$, $\sigma = 1$, and $\alpha = -0.2$, $\eta = 1$, $a = 1.5$, $b = 80$.

Table 7: Simulation results for the estimator of the mean quality adjusted lifetime, $\hat{\mu}$, and its asymptotic standard error for 40% censored data based on 500 replications.

True model covariate distribution	True μ	N	Baseline covariate (Fit)	Bias($\hat{\mu}$)	SD($\hat{\mu}$)	MeanSE	95% CP
$V_1 \sim BIN(0.5)$	8.577	200	Known β	0.008	0.109	0.111	0.952
			Binary	0.009	0.108	0.113	0.954
			Continuous	0.016	0.111	0.114	0.948
		400	Known β	0.006	0.080	0.079	0.940
			Binary	0.006	0.079	0.081	0.950
			Continuous	0.016	0.082	0.084	0.944
		800	Known β	0.002	0.056	0.056	0.948
			Binary	0.002	0.056	0.059	0.960
			Continuous	0.011	0.061	0.064	0.956
$V_2 \sim N(0,1)$	8.980	200	Known β	0.009	0.123	0.124	0.948
			Continuous	0.009	0.123	0.126	0.954
			Binary	0.018	0.125	0.127	0.951
		400	Known β	0.009	0.088	0.088	0.944
			Continuous	0.009	0.087	0.091	0.956
			Binary	0.014	0.092	0.094	0.952
		800	Known β	0.001	0.061	0.062	0.958
			Continuous	0.001	0.060	0.066	0.972
			Binary	0.007	0.065	0.064	0.964
V_1 and $V2$	8.780	200	Known β	<0.001	0.058	0.058	0.956
			Binary and Continuous	<0.001	0.059	0.061	0.962
			Binary	0.015	0.062	0.064	0.958
			Continuous	0.009	0.061	0.063	0.959
		400	Known β	0.001	0.043	0.041	0.944
			Binary and Continuous	0.001	0.044	0.044	0.954
			Binary	0.013	0.053	0.049	0.948
			Continuous	0.008	0.048	0.047	0.952
		800	Known β	0.001	0.032	0.029	0.939
			Binary and Continuous	0.001	0.033	0.034	0.960
			Binary	0.008	0.038	0.036	0.952
			Continuous	0.003	0.035	0.036	0.956

$\beta_0 = 2$, $\beta_1 = -3$, $\sigma_{11} = 1$, $\sigma_{12} = \sigma_{21} = 0$, $\sigma_{22} = 0.5$, $\sigma = 1$, and $\alpha = -0.2$, $\eta = 1$, $a = 1.5$, $b = 80$.

BIBLIOGRAPHY

[1] N. K. Aaronson, S. Ahmedzai, B. Bergman, M. Bullinger, A. Cull, N. J. Duez, A. Filiberti, H. Flechtner, S. B. Fleishman, J. C. J. M. Haes, S. Kaasa, M. Klee, D. Osoba, D. Razavi, P. B. Rofe, S. Schraub, K. Sneeuw, M. Sullivan, and F. Takeda. The european organization for research and treatment of cancer qlq-c30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 85:365–376, 1993.

[2] A. Andrei and S. Murray. Regression models for the mean of the quality-of-life-adjusted restricted survival time using pseudo-observations. *Biometrics*, 63:398–404, 2007.

[3] W. E. Barlow. Robust variance estimation for the case-cohort design. *Biometrics*, 50:1064–1072, 1994.

[4] O. Borgan, B. Langholz, S. O. Samuelsen, L. Goldstein, and J. Pogoda. Exposure stratified case-cohort designs. *Lifetime Data Analysis*, 6:39–58, 2000.

[5] H. Y. Chen. Fitting semiparametric transformation regression models to data from a modified case-cohort design. *Biometrika*, 88:255–268, 2001.

[6] K. Chen and S. H. Lo. Case-cohort and case-control analysis with cox's model. *Biometrika*, 86:755–764, 1999.

[7] K. N. Chen. Generalized case-cohort sampling. *Journal of the Royal Statistical Society*, 62:449–460, 2001.

[8] B. F. Cole, R. D. Celber, and A. Goldhirsch. Cox regression models for quality adjusted survival analysis. *Statistics in Medicine*, 12:975–987, 1993.

[9] D. R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society*, 34:187–220, 1972.

[10] K. J. Cullen and C. A. P. Boundy. Factors relating to behaviour disorders in children. *Journal of Paediatrics and Child Health*, 2:70–80, 1966.

[11] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:1–38, 1977.

[12] C. L. Faucett and D. C. Thomas. Simultaneously modelling censored survival data and repeatedly measured covariates: a gibbs sampling approach. *Statistics in Medicine*, 15:1663–1685, 1996.

[13] T. R. Fleming and D. P. Harrington. *Counting Process and Survival Analysis*. New York: Wiley, 1991.

[14] R. D. Gelber, R. S. Gelman, and A. Goldhirsch. A quality of life oriented endpoint for comparing therapies. *Biometrics*, 45:781–795, 1989.

[15] R. D. Gelber, A. Goldhirsch, and F. Cavalli. Quality-of-life-adjusted evaluation of a randomized trial comparing adjuvant therapies for operable breast cancer (for the international breast cancer study group). *Annals of Internal Medicine*, 114:621–628, 1991.

[16] R. D. Gill. *Censoring and stochastic integrals, Mathematical Centre Tracts No. 124.* Amsterdam; Mathematisch Centrum, 1980.

[17] P. P. Glasziou, R. J. Simes, and R. D. Gelber. Quality adjusted survival analysis. *Statistics in Medicine*, 9:1259–1276, 1990.

[18] A. Goldhirsch, R. D. Gelber, R. J. Simes, P. P. Glasziou, and A Coates for the Ludwig Breast Cancer Study Group. Costs and benefits fo adjuvant therapy in breast cancer: a quality-adjusted survival analysis. *Journal of Clinical Oncology*, 7:36–44, 1989.

[19] R. Henderson, P. Diggle, and A. Dobson. Joint modelling of longitudinal measurements and event time data. *Biometrics*, 1:465–480, 2000.

[20] J. W. Hogan and N. M. Laird. Mixture models for the joint distribution of repeated measures and event times. *Statistics in Medicine*, 16:239–257, 1997.

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

[22] F. Hsieh, Y. K. Tseng, and J. L. Wang. Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics*, 62:1037–1043, 2006.

[23] Y. Huang and T. A. Louis. Expressing estimators of expected quality adjusted survival as functions of nelson-aalen estimators. *Lifetime data analysis*, 5.3:199–212, 1999.

[24] J. S. Hwang, J. Y. Tsauo, and J. D. Wang. Estimation of expected quality-adjusted survival by cross-sectional survey. *Statistics in Medicine*, 15:93–102, 1996.

[25] J. A. Kellum, L. Kong, M. P. Fink, L. A. Weissfeld, D. M. Yealy, M. R. Pinsky, J. Fine, A. Krichevsky, R. L. Delude, and D. C. Angus. Understanding the inflammatory cytokine response in pneumonia and sepsis. *Archives of Internal Medicine*, 167:1655–1663, 2007.

[26] L. Kong and J. Cai. Case-cohort analysis with accelerated failure time model. *Biometrics*, 65:135–142, 2009.

[27] L. Kong, J. Cai, and P. K. Sen. Weighted estimating equations for semiparametric transformation models with censored data from a case-cohort design. *Biometrika*, 91:305–319, 2004.

[28] S. Kong and J. Cai. Marginal hazards model for case-cohort studies with multiple disease outcomes. *Biometrika*, 96.4:887–901, 2009.

[29] E. L. Korn. On estimating the distribution function for quality of life in cancer clinical trials. *Biometrika*, 80:535–542, 1993.

[30] M. Kulich and D. Y. Lin. Additive hazards regression for case-cohort studies. *Biometrika*, 87:73–87, 2000.

[31] B. Langholz and J. Jiao. Computational methods for case-cohort studies. *Computational Statistics & Data Analysis*, 51:3737–3748, 2007.

[32] D. Y. Lin and Z. Ying. Cox regression with incomplete covariate measurements. *Journal of the American Statistical Association*, 88:1341–1349, 1993.

[33] R. J. A. Little and D. B. Rubin. *Statistical analysis with missing data (2nd ed.)*. New York: Wiley, 2002.

[34] L. Liu and X. Huang. The use of gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine*, 27:2665–2683, 2008.

[35] W. Lu and A. A. Tsiatis. Semiparametric transformation models for the case-cohort study. *Biometrika*, 93:207–214, 2006.

[36] R. H. Lyles, C. M. Lyles, and D. J. Taylor. Random regression models for human immunodeficiency virus ribonucleic acid data subject to left censoring and informative drop-outs. *Applied Statistics*, 49:485–497, 2000.

[37] S. Ma. Additive risk model with case-cohort sampled current status data. *Statistical Papers*, 48:595–608, 2007.

[38] B. Nan, M. Yu, and J. D. Kalbfleiisch. Censored linear regression for case-cohort studies. *Biometrika*, 93:747–762, 2006.

[39] B. Pradhan and A. Dewanji. Parametric estimation of quality adjusted lifetime (qal) distribution in progressive illness-death model. *Statistics in Medicine*, 28:2012–2027, 2009.

[40] R. L. Prentice. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika*, 73:1–11, 1986.

[41] J. M. Robins and A. Rotnitzky. Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiology-Methodological Issues*, pages 297–331, 1992.

[42] J. M. Robins, A. Rotnitzky, and L. P. Zhao. Estimation of regression coefficients when some regressors are not always observed. *Journal of American Statistics Association*, 89:846–866, 1994.

[43] A. Rotnitzky and J. M. Robins. Inverse probability weighting in survival analysis. *Encyclopedia of Biostatistics*, 2005.

[44] S. O. Samuelsen, H. Anestad, and A. Skrondal. Stratified case-cohort analysis of general cohort sampling designs. *Scandinavian Journal of Statistics*, 34:103–119, 2007.

[45] S. G. Self and R. L. Prentice. Asymptotic distribution theory and efficiency results for case-cohort studies. *The Annals of Statistics*, 16:64–81, 1988.

[46] W. J. Shih. Problems in dealing with missing data and informative censoring in clinical trials. *Trials*, 3.1:4, 2002.

[47] P. Sorensen and P. K. Anderson. Competing risks analysis of the case-cohort design. *Biometrika*, 87:49–59, 2000.

[48] J. Sun, L. Sun, and N. Flournoy. Additive hazards model for competing risks analysis of the case-cohort design. *Communications in Statistics: Theory and Methods*, 33:351–366, 2004.

[49] The ARIC Investigators. The atherosclerosis risk in communities (aric) study: design and objectives. *American Journal of Epidemiology*, 129:687–702, 1989.

[50] T. M. Therneau and H. Li. Computing the cox model for case cohort designs. *Lifetime Data Analysis*, 5:99–112, 1999.

[51] C. Tseng and M. Liu. Joint modeling of survival data and longitudinal measurements under nested case-control sampling. *Statistics in Biopharmaceutical Research*, 1:415–423, 2009.

[52] Y. K. Tseng, F. Hsieh, and J. L. Wang. Joint modeling of accelerated failure time and longitudinal data. *Biometrika*, 92:587–603, 2005.

[53] M. J. van de Laan and A. Hubbard. Locally efficient estimation of the quality-adjusted lifetime distribution with right-censored data and covariates. *Biometrics*, 55:530–536, 1999.

[54] V. Velanovich, S. R. Vallance, J. R. Gusz, F. V. Tapia, and M. A. Harkabus. Quality of life scale for gastroesophageal reflux disease. *Journal of the American College of Surgeons*, 183:217–224, 1996.

[55] E. F. Vonesh, T. Greene, and M. D. Schluchter. Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine*, 25:143–163, 2006.

[56] S. Wacholder, M. H. Gail, D. Pee, and R. Brookmeyer. Alternative variance and efficiency calculations for the case-cohort design. *Biometrika*, 76:117–123, 1989.

[57] H. Wang and H. Zhao. Regression analysis of mean quality-adjusted lifetime with censored data. *Biostatistics*, 8:368–382, 2007.

[58] M. S. Wulfsohn and A. A. Tsiatis. A joint model for survival and longitudinal data measured with error. *Biometrics*, 53:330–339, 1997.

[59] J. Xie and C. Liu. Adjusted kaplan-meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Statistics in Medicine*, 24:3089–3110, 2005.

[60] D. Zeng, D. Y. Lin, C. L. Avery, K. E. North, and M. S. Bray. Efficient semiparametric estimation of haplotype-disease associations in case-cohort and nested case-control studies. *Biometrics*, 7:486–502, 2006.

[61] H. Zhang, D. E. Schaubel, and J. D. Kalbfleisch. Proportional hazards regression for the analysis of clustered survival data from case-cohort studies. *Biometrics*, 67:18–28, 2011.

[62] H. Zhao and A. A. Tsiatis. A consistent estimator for the distribution of quality of adjusted survival time. *Biometrika*, 84:339–348, 1997.

[63] H. Zhao and A. A. Tsiatis. Estimating mean quality adjusted lifetime with censored data. *Biostatistics*, 62:175–188, 2000.

[64] Y. Zhao and H. Wang. Empirical likelihood inference for the regression model of mean quality-adjusted lifetime with censored data. *The Canadian Journal of Statistics*, 36:463–478, 2008.