Spring 5-3-2011

Treatment Comparison in Biomedical Studies Using Survival Function

Meng Zhao

TREATMENT COMPARISON IN BIOMEDICAL STUDIES USING SURVIVAL FUNCTION

by

MENG ZHAO

Under the Direction of Dr. Yichuan Zhao

ABSTRACT

In the dissertation, we study the statistical evaluation of treatment comparisons by evaluating the relative comparison of survival experiences between two treatment groups. We construct confidence interval and simultaneous confidence bands for the ratio and odds ratio of two survival functions through both parametric and nonparametric approaches.

We first construct empirical likelihood confidence interval and simultaneous confidence bands for the odds ratio of two survival functions to address small sample efficacy and sufficiency. The empirical log-likelihood ratio is developed, and the corresponding asymptotic distribution is derived. Simulation studies show that the proposed empirical likelihood band has outperformed the normal approximation band in small sample size cases in the sense that it yields closer coverage probabilities to chosen nominal levels.

Furthermore, in order to incorporate prognostic factors for the adjustment of survival functions in the comparison, we construct simultaneous confidence bands for the ratio and odds ratio of survival functions based on both the Cox model and the additive risk model. We develop simultaneous confidence bands by approximating the limiting distribution of cumulative hazard functions by zero-mean Gaussian processes whose distributions can be generated through Monte Carlo simulations. Simulation studies are conducted to evaluate the performance for proposed models. Real applications on published clinical trial data sets are also studied for further illustration purposes.

In the end, the population attributable fraction function is studied to measure the impact of risk factors on disease incidence in the population. We develop semiparametric estimation of attributable fraction functions for cohort studies with potentially censored event time under the additive risk model.

INDEX WORDS: Additive risk model, Attributive fraction function, Censoring, Confidence band, Counting Process, Cox Regression, Empirical likelihood, Martingale, Odds ratio, Ratio, Right censored data, Survival function

TREATMENT COMPARISON IN BIOMEDICAL STUDIES USING SURVIVAL FUNCTION

by

MENG ZHAO

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2011

TREATMENT COMPARISON IN BIOMEDICAL STUDIES USING SURVIVAL FUNCTION

by

MENG ZHAO

Committee Chair: Dr. Yichuan Zhao

Committee: Dr. Yuanhui Xiao

Dr. Xu Zhang

Dr. Yanqing Zhang

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2011

DEDICATION

To my parents and my advisor.

ACKNOWLEDGEMENTS

In writing this dissertation, I have received substantial help from a great variety of people with warm patience and politeness. I am not able to thank everyone by name, but I would love to record my great debts of gratitude to the following people.

First and foremost, my gratitude to my advisor, Dr. Yichuan Zhao, is so unspeakable that it becomes so difficult for me to find a word from dictionary to sufficiently express how grateful I am to my advisor for all of his support, supervision, patience and understanding. Dr. Zhao has suffered and sacrificed quite much for cultivating me to be another Dr. Zhao. There is no way I could have finished this dissertation research without his provision. It is the kind of help I have always rested assured with and relied on. My mind has truly been opened to the splendid world of academic research by Dr. Zhao's generous mentorship for more than four years. He has pushed me continuously which guided me almost step by step throughout my entire adventure of pursuing a Ph.D. degree.

I would also like to take this great opportunity to acknowledge other members of my dissertation committee, Dr. Yuanhui Xiao, Dr. Xu Zhang and Dr. Yanqing Zhang, for taking their time to read this dissertation and provide instructive comments, especially Dr. Xu Zhang for checking my dissertation on a verbal basis that fully restructured the contextual organization.

Special thanks to all the faculty and staff in the Department of Mathematics and Statistics in Georgia State University for their continuous help and kindness. It would not have been possible for me to complete the requirements of the Ph.D. program without their guidance and support.

My friends Haochuan Zhou, Xin Huang, Binhuan Wang, Shuang Liu and Dirk Gilmore, among others, have been enthusiasm personified throughout for their suggestions, discussions and criticisms. They have always been there for me with endless encouragement and wisdom from many aspects. This is the treasure worthwhile for a life-long cherishing.

I thank Professor Mei-Jie Zhang, the IBMTR and the German CML Study Group for providing us with the CML data, and Professor Song Yang for providing us with the gastric cancer data. I also thank Dr. Ian McKeauge for his enlightening comments that motivate the first half of my dissertation.

Above all, I give the deepest and most heartfelt gratitude to my parents for giving me everything that they have. I hope I could have the opportunity to give mine in return.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

In clinical trials and related medical studies, physicians and medical researchers often impose significant concentrations on the comparison of two treatments in order to justify the effect of a new medicine or a new cure. Statistical analysis usually provides important reference to the quantitative evaluation of medical advantages of one treatment over another. Modeling inaccuracy, however, might sometimes result in severe damages both financially and ethically, especially to those studies involving human beings. Such concern has triggered my research interest to focus on the development of statistical methods on the evaluation of treatment comparison results.

We study both parametric and nonparametric statistical approaches for constructing confidence intervals and simultaneous confidence bands for the ratio and odds ratio of survival functions of two targeted groups. We develop methods to account for adjusted covariates in order to improve the practical applicability of our models. Methodological details are presented in the following, along with future research plans.

Due to the well-known small sample efficacy of empirical likelihood based methods, we first adopt the empirical likelihood for constructing simultaneous confidence bands for the odds ratio of two survival functions derived from the type of clinical trial data that has two groups receiving different treatments. For more practical convenience, the empirical likelihood confidence interval is also given. The empirical log-likelihood ratio function is developed, and the corresponding asymptotic distribution is derived. Moreover, the conventional normal approximation method is also implemented in comparison with the empirical likelihood based method in order to demonstrate the methodological advantage of our proposed method. Simulation studies show that the proposed empirical likelihood band has outperformed the normal approximation band in small sample size cases in the sense that it yields closer coverage probabilities to chosen nominal levels.

When diagnostic results indicate that a certain treatment has time-varying effects, in order to better assess a treatment over another, it becomes both clinically meaningful and statistically reasonable to incorporate information from prognostic factors to allow necessary adjustments of

survival functions. For such reason, we developed methods to construct simultaneous confidence bands for both the ratio and odds ratio of two survival functions under both the multiplicative risk model and the additive risk model to provide physicians with more information for diagnostic judgment.

Based on the martingale central limit theorem, we can show that both the process of survival ratio and survival odds ratio converge weakly to zero-mean Gaussian processes (Andersen et al. [5]). It is commonly recognized that the limiting Gaussian process does not have independent increment, and hence numerical realization of the limiting distribution becomes quite intricate, which hinders the construction of confidence bands. Lin et al. [47] propose a simulation method in resolving such issues for a single survival function. Our approach in this dissertation research uses a similar method by uniformly consider both samples under a stratified Cox's regression. Simulation results have shown coverage probabilities of proposed confidence bands to be sufficiently accurate toward to the normal level. Further illustrations are fulfilled by two real clinical trial data sets.

Equivalently important as the multiplicative risk model, the additive risk model is another principle framework for evaluating the association between risk factors and survival. Such model possesses great analytical value for the statistical evaluation of treatment comparison. Most importantly, systemized by Lin and Ying [51], the additive model has a closed form for the estimator of regression parameters so that the two-sample formulation becomes much more straightforward with significantly reduced computational cost. Song et al. [74] and Yin and Hu [81] have studied confidence bands for the survival function as well as the cumulative hazard function based on the additive risk model. Lee and Hyun [42] extended the work to study the difference of two survival functions. But to the best of our knowledge, confidence bands that accompany the ratio or odds ratio of two survival functions remain unavailable. Since the ratio of two survival functions can be treated as an indicator of the relative risk, and the odds ratio carries the idea of odds that has many special applications in epidemiology, it is better to consider both the ratio and the odds ratio together for complete methodological reference. We developed simultaneous confidence bands for such purpose, designed for a variety of specific applications. Martingale processes under the additive risk model fail to retain the independent increment structure, and therefore cannot be transformed into a standard Brownian bridge. Asymptotically equivalent processes have to be considered analo-

gously to those mentioned by Lin et al. [47] in order to search for proper critical values to construct simultaneous confidence bands.

Chapter 2

EMPIRICAL LIKELIHOOD FOR THE ODDS RATIO OF TWO SURVIVAL FUNCTIONS WITH RIGHT CENSORING

In the present chapter, we study pointwise confidence intervals as well as simultaneous confidence bands for the odds ratio of two survival functions, by introducing the empirical likelihood (EL) method. We develop EL procedures for two-sample survival odd ratios and establish major theoretical results. Simulation studies are conducted to compare the relative performance between the proposed method and the normal approximation method. Proofs are relegated to the Appendix A.

2.1 Motivation

Being an alternative quantity in the measure of association, the odds ratio can be obtained easily from either a cohort study or a clinical trial. Concentrations on the odds ratio have become popular in recent literature of biomedical researches. More and more physicians and epidemiologists began to adopt the odds ratio in analyzing clinical trial results and evaluating epidemiological findings. See Zhang and Yu [85] for a discussion about the odds ratio in cohort studies and Cummings [20] for a famous example.

Recall bias is a big issue in case-control studies. Consider, for instance, a case-control study on individuals over 80 years of age in terms of early-stage Alzheimer's disease, where cases are those with the disease and controls are those showing no symptoms. The time-to-event, a fairly difficulty event for elderly people to recall, is the time elapsed since last reading a newspaper (See Hassan [32] for a recent discussion of the general issue of recall bias). We believe that under time reversal, one way of resolving the recall bias issue is to consider the data with recall bias as a right censored data with lack of recall beyond certain time in the past causing the right censoring. The time-to-event can be modeled by survival functions, therefore comparing the odds of survival functions between cases and controls will furnish a mathematically reasonable measurement in evaluating cases versus controls.

Furthermore, it is both statistically meaningful and clinically referential to construct confidence bands for the odds ratio of survival functions as a graphical test for the proportional odds model. For instance, to see if there is any diagnostic advantage of one treatment over another, one could check whether the constructed confidence band contains an indicator line of constant one. There is a formal test for the proportional odds (Dauxois and Kirmani [23]), but a graphical test might be more enlightening and intuitive.

The normal approximation (NA) based confidence regions have some known limitations. For instance, confidence intervals constructed are always symmetric, which may not be desirable in every situation, and the coverage probability of a $100(1-\alpha)\%$ interval estimator for the true survival odds ratio is also noticeably less than the nominal level when the sample size is small (as shown in our simulation study). Compared to a normal approximation one, EL-based confidence regions have the following advantages: (1) It reduces complexity on statistical inferences, due to the dispensability of deriving a variance estimator; (2) It produces better coverage accuracy in small sample cases; (3) It is not necessarily symmetric, which allows the resulting interval and band estimates to better reflect the shape of a specific data. Moreover, EL-based confidence regions are range-preserving, transform-respecting and asymmetric since they rely solely on the features of the data to determine the shape of the confidence region.

These considerations motivate us to propose an EL approach for constructing simultaneous confidence bands for the odds ratio of two survival functions.

Many authors have investigated the use of the EL approaches in survival analysis. The first contribution is attributed to Thomas and Grunkemeier [75]. Subsequently, Owen [57, 58] introduced the EL terminology and proved a number of fundamental results. Owen [61] further provides a comprehensive summary of EL methods. Li [44] and Murphy [57] developed theoretical foundations for deriving empirical likelihood ratio functions with censored data. Li et al. [45] review and summarize the various results of many literatures using EL in time-to-event problems. Particularly, McKeague and Zhao [55] constructed a simultaneous confidence band for the ratio of two survival functions based on independent, right-censored data. In the subsequent work, McKeague and Zhao [56] develop a method of estimating both the difference and the ratio of two distribution functions based on the EL method. Their method can be extended to estimate either the difference or ratio of the unadjusted cumulative hazard functions. Shen and He [71] derived EL based confidence bands

for the difference of two survival functions. Ren [66] applied weighted EL to case-control logistic regression models.

2.2 Main Results

We consider two independent samples with right censoring. For $j = 1, 2$, let $\{T_{ji}, i = 1, 2, \ldots, n_j\}$ be i.i.d. failure times from the distribution F_j. Let $\{C_{ji}, i = 1, 2, \ldots, n_j\}$, from the distribution G_j, be non-negative i.i.d. censoring times independent of failure times. With right censoring, we denote observations for each sample as

$$(X_{ji}, \delta_{ji}), \tag{2.1}$$

where $X_{ji} = \min(T_{ji}, C_{ji})$, and $\delta_{ji} = I(T_{ji} \leq C_{ji})$ is the censoring indicator. Note that throughout the chapter, we use $I(A)$ for the indicator function of set A. Define the odds ratio of two survival functions as

$$\theta(t) = \frac{1 - S_1(t)}{S_1(t)} \Big/ \frac{1 - S_2(t)}{S_2(t)} .$$

With the loss of generality, let $0 \leq T_{j1} \leq \cdots T_{jN_j} < \infty$ be ordered uncensored survival times of sample $j = 1, 2$, and write

$$r_{ji} = \sum_{k=1}^{n_j} I(X_{jk} \geq T_{ji}), \quad d_{ji} = \sum_{k=1}^{n_j} I(X_{ji} = T_{ji}, \delta_{jk} = 1)$$

to denote the number of subjects "at risk" prior to time T_{ji} and "dead" at time T_{ji}, respectively. Moreover, define

$$K_j(t) = \sum_{i=1}^{N_j} I(T_{ji} \leq t), \quad j = 1, 2.$$

Let Γ be the space of all survival functions defined on $[0, \infty)$. For any $S_1, S_2 \in \Gamma$, the empirical likelihood function is defined as

$$L(S_1, S_2) = \prod_{j=1}^{2} \prod_{i=1}^{n_j} \left[S_j(X_{ji}-) - S_j(X_{ji})\right]^{\delta_{ji}} \left[S_j(X_{ji})\right]^{1-\delta_{ji}} .$$

For fixed t, denote the true odds ratio as $\theta_0(t) = \frac{1-S_1(t)}{S_1(t)} \Big/ \frac{1-S_2(t)}{S_2(t)} > 0$. The empirical likelihood ratio can be written as

$$\mathcal{R}(\theta_0, \eta, t) = \frac{\sup\left\{L(S_1, S_2) : S_1(t) = \frac{\eta}{\eta+\theta_0(t)-\theta_0(t)\eta}, S_2(t) = \eta, S_1, S_2 \in \Gamma\right\}}{\sup\{L(S_1, S_2) : S_1, S_2 \in \Gamma\}}$$

where η is a nuisance parameter such that $0 < \eta < 1$.

Referring to Li [44], we can rewrite the log-likelihood ratio function as:

$$\ln \mathcal{R}(\theta_0, \eta, t) = \sum_{j=1}^{2} \sum_{i=1}^{K_j(t)} \left\{(r_{ji} - d_{ji}) \ln\left(1 + \frac{\lambda_j}{r_{ji} - d_{ji}}\right) - r_{ji} \ln\left(1 + \frac{\lambda_j}{r_{ji}}\right)\right\}, \tag{2.2}$$

where the Lagrange multipliers $\lambda_1,\ \lambda_2$ satisfy

$$\sum_{i=1}^{K_1(t)} \ln\left(1 - \frac{d_{1i}}{r_{1i} + \lambda_1}\right) - \ln\left(\frac{\eta}{\eta + \theta_0(t) - \theta_0(t)\eta}\right) = 0, \tag{2.3}$$

$$\sum_{i=1}^{K_2(t)} \ln\left(1 - \frac{d_{2i}}{r_{2i} + \lambda_2}\right) - \ln(\eta) = 0. \tag{2.4}$$

Furthermore, to maximize $\ln \mathcal{R}(\theta_0, \eta, t)$, η satisfies

$$\lambda_1 \frac{\partial E_1(\eta, \lambda_1, \lambda_2, t)}{\eta} + \lambda_2 \frac{\partial E_2(\eta, \lambda_1, \lambda_2, t)}{\eta} = 0,$$

where, $E_1(\eta, \lambda_1, \lambda_2, t)$ and $E_2(\eta, \lambda_1, \lambda_2, t)$ denote the left hand sides of equation (2.3) and (2.4), respectively. It follows that,

$$\frac{\lambda_1}{n}\left(\frac{\theta_0(t)}{\eta + \theta_0(t) - \eta\theta_0(t)}\right) + \frac{\lambda_2}{n} = 0. \tag{2.5}$$

In order to present our main result, we introduce the following notations. For any cdf F, denote $\bar{F} = 1 - F$ and define a_F, b_F of F as

$$a_F = \inf\{x : F(x) > 0\} \quad \text{and} \quad b_F = \sup\{x : F(x) < 1\}.$$

Additionally, define

$$\sigma_j^2(t) = \int_0^t \frac{\mathrm{d}F_j(s)}{S_j(s-)\bar{H}_j(s)}, \quad j = 1, 2, \tag{2.6}$$

where for sample j, F_j is the distribution function of the survival time X_{ji} and H_j is that of the censored time Z_{ji}.

Write

$$\sigma^2(t) = \frac{(1 - S_2)^2\sigma_1^2(t)}{p_1} + \frac{(1 - S_1)^2\sigma_2^2(t)}{p_2}, \tag{2.7}$$

where p_1 and p_2 are constant numbers between 0 and 1.

Theorem 1. *Let $\tau_1, \tau_2 \in \mathbb{R}$ such that $a_{F_1} \vee a_{F_2} < \tau_1 < \tau_2 < b_{H_1} \wedge b_{H_2}$ and suppose $n_j/n \to p_j \in (0, 1)$, as $n \to \infty$, where $n = n_1 + n_2$, then for any $t \in [\tau_1, \tau_2]$, there exists a solution $\eta_E(t)$ to equation (2.5) almost surely, such that $\eta_E(t) = \arg\sup_\eta \mathcal{R}(\theta_0, \eta, t)$, and*

$$-2\log\mathcal{R}(\theta_0, \eta, \cdot) \xrightarrow{\mathfrak{D}} \frac{W^2(\cdot)}{\sigma^2(\cdot)}, \tag{2.8}$$

where

$$W(t) = \frac{(1 - S_2(t))W_1\left(\sigma_1^2(t)\right)}{\sqrt{p_1}} + \frac{(1 - S_1(t))W_2\left(\sigma_2^2(t)\right)}{\sqrt{p_2}}, \tag{2.9}$$

and W_1, W_2 are independent standard Brownian motions.

Corollary 1. *Assuming that $n_j/n \to p_j$ as $n \to \infty$, then for any t such that $a_{F_1} \vee a_{F_2} < t < b_{H_1} \wedge b_{H_2}$, we have*

$$-2\log\mathcal{R}(\theta_0, \eta, t) \xrightarrow{\mathfrak{D}} \chi_1^2 \ . \tag{2.10}$$

Now, we will show how to construct the empirical likelihood confidence band for $\theta_0(t)$. By Theorem 1 and the continuous mapping theorem, we can show that

$$\sup_{t\in[\tau_1,\tau_2]} \{-2\sigma^2(t)\log\mathcal{R}(\theta_0, \eta, t)\} \xrightarrow{\mathfrak{D}} \sup_{t\in[\tau_1,\tau_2]} W^2(t).$$

From Andersen et al. [5], p. 262, we know

$$\hat{\sigma}_j^2(t) = n_j \sum_{i=1}^{K_j(t)} \frac{d_{ji}}{r_{ji}(r_{ji} - d_{ji})} = \sigma_j^2(t) + o_p(1), \quad \text{as } n_j \to \infty, \quad j = 1, 2, \tag{2.11}$$

which yield that

$$\hat{\sigma}^2(t) = \frac{(1 - \hat{S}_2)^2 \hat{\sigma}_1^2(t)}{\hat{p}_1} + \frac{(1 - \hat{S}_1)^2 \hat{\sigma}_2^2(t)}{\hat{p}_2} \tag{2.12}$$

is a consistent estimator for $\sigma^2(t)$, where $\hat{S}_j(t)$ is the Kaplan-Meier estimator of $S_j(t)$.

Thus, we can construct the asymptotic $100(1-\alpha)\%$ confidence band for $\theta_0(t)$ as

$$I_{n,\alpha} = \{(t, \theta) : -2\hat{\sigma}^2(t) \log \mathcal{R}(\theta, \eta_E, t) \leq K_\alpha^2[\tau_1, \tau_2], t \in [\tau_1, \tau_2]\}, \tag{2.13}$$

where $K_\alpha[\tau_1, \tau_2]$ is the upper α-quantile of the distribution of $\sup_{t \in [\tau_1, \tau_2]} |W(t)|$. Practically, based on Lin et al. [47], Monte Carlo methods can be used to simulate such distribution.

For fixed t, by Corollary 1, a pointwise confidence interval with asymptotic coverage probability of $1 - \alpha$ is given as

$$I_{n,\alpha}(t) = \{\theta : -2 \log \mathcal{R}(\theta, \eta_E, t) \leq C_\alpha\}$$

where C_α is the upper α-quantile of χ_1^2.

2.3 Simulation Study

We conduct a simulation study to compare the performance of the proposed EL method and a normal approximation (NA) based method in terms of coverage probability of both pointwise confidence intervals and simultaneous confidence bands.

We first derive the NA type confidence band for $\theta_0(t)$. Write

$$\hat{\theta}_n(t) = \frac{1 - \hat{S}_1(t)}{\hat{S}_1(t)} \Big/ \frac{1 - \hat{S}_2(t)}{\hat{S}_2(t)},$$

where $\hat{S}_1(t)$ and $\hat{S}_2(t)$ are Kaplan-Meier estimators for S_1 and S_2. Note that according to the CLT of $\hat{S}_j(t)$ (Andersen et al. [5]), we know

$$n_j^{1/2}(\hat{S}_j(\cdot) - S_j(\cdot)) \xrightarrow{\mathfrak{D}} S_j(\cdot)W_j(\sigma_j^2(\cdot)), \quad j = 1, 2.$$

Thus, it follows by the functional delta method that

$$\sqrt{n}\left(\hat{\theta}(\cdot) - \theta(\cdot)\right) \xrightarrow{\mathscr{D}} W^*(\cdot), \tag{2.14}$$

where

$$W^*(\cdot) = \frac{S_2(\cdot)}{S_1(\cdot)(1 - S_2(\cdot))^2} \cdot \left[\frac{(1 - S_2(\cdot))W_1(\sigma_1^2(\cdot))}{\sqrt{p_1}} + \frac{(1 - S_1(\cdot))W_2(\sigma_2^2(\cdot))}{\sqrt{p_2}}\right].$$

Therefore, the asymptotic $100(1-\alpha)\%$ confidence interval of $\theta_0(t)$ is

$$I^*_{n,\alpha}(t) = \hat{\theta}_n(t) \pm z_{\alpha/2}\frac{\hat{S}_2(t)}{\hat{S}_1(t)(1 - \hat{S}_2(t))^2}\frac{\hat{\sigma}^*(t)}{\sqrt{n}},$$

where

$$\hat{\sigma}^{*2}(t) = \frac{(1 - \hat{S}_2(t))^2\hat{\sigma}_1^2(t)}{\hat{p_1}} + \frac{(1 - \hat{S}_1(t))^2\hat{\sigma}_2^2(t)}{\hat{p_2}}.$$

Similarly, the asymptotic $100(1-\alpha)\%$ normal approximation type confidence band for $t \in [\tau_1, \tau_2]$ is defined as

$$\hat{\theta}_n(t) \pm n^{-1/2}K^*_\alpha[\tau_1, \tau_2], \quad t \in [\tau_1, \tau_2], \tag{2.15}$$

where $K^*_\alpha[\tau_1, \tau_2]$ is the upper α-quantile of the distribution of $\sup_{t\in[\tau_1,\tau_2]} |W^*(t)|$, which can be generated through Monte Carlo methods.

In our simulations, both survival time and censoring time are generated independently from Exponential distributions. Four different sample sizes, i.e., 30, 50, 80 and 100, respectively, have been selected for the first sample and identical sample sizes have been assigned to the second one correspondingly. In addition, two different censoring rates (CR), 10% and 30%, are selected. Under each setting, 1000 repetitions are conducted to measure coverage probabilities through Monte Carlo simulations. More specifically, we set $F_1 = \text{Exp}(2.3)$, $F_2 = \text{Exp}(1)$ and generate censoring

distributions $G_1 = \text{Exp}(20)$, $G_2 = \text{Exp}(8)$ to generate a 10% CR, while $G_1 = \text{Exp}(9)$, $G_2 = \text{Exp}(2.3)$ to guarantee a 30% CR.

Two different nominal levels, $\alpha = 0.90$ and $\alpha = 0.95$ are chosen to compare the performance of our empirical likelihood (EL) confidence band with that of the normal approximation (NA) type one. We set $\tau_1 = 0.1$ and $\tau_2 = 2.5$ to generate $K_\alpha[\tau_1, \tau_2]$ and $K^*_\alpha[\tau_1, \tau_2]$. When sample size is small, τ_2 will be adjusted to guarantee an effective sample size, $\#\{(i,j) : T_{ji} >= \tau_2\}$, at τ_2 of at least 10% of the total sample size in avoiding instability (Hollander et al. [33]). Detailed results are reported in Table 1.

For fixed t, our proposed empirical likelihood confidence interval, $I_{n,\alpha}(t)$ and the normal approximation one, $I^*_{n,\alpha}(t)$, are also studied for further justification. Simulation results at $t = 1.1$ are summarized in Table 2 in Appendix B.

From the above two tables, we may be able to make the following conclusions. It can shown that both EL confidence bands and confidence intervals generally give more accurate coverage probabilities in small sample sizes, which are much closer to the nominal level than those of the normal approximation method. More specifically, simultaneous empirical likelihood confidence bands generally outperform the normal approximation type bands for virtually all different sample sizes, with more stable and consistent performance. As expected, the advantages of the proposed empirical likelihood methods for constructing pointwise confidence intervals gradually disappear as sample size becomes larger.

2.4 Remarks

The aforementioned method can be further extended to semiparametric estimators with adjusted covariates. Consider

$$\theta(t|Z_1, Z_2) = \frac{1 - \hat{S}_1(t|Z_1)}{\hat{S}_1(t|Z_1)} \Big/ \frac{1 - \hat{S}_2(t|Z_2)}{\hat{S}_2(t|Z_2)},$$

where Z_1 are possible time-varying covariates of sample 1, and Z_2 are those of sample 2.

Incorporating adjusted covariates to empirical likelihood based estimator will provide useful information of more relevant risk factors while at the same time maintaining small sample efficiencies of the empirical likelihood based method. Moreover in many applications, it is more reasonable to

assume the survival and censoring time to be independent depending on covariates (risk factors), then introducing a regression model, say, the Cox model, will become quite necessary.

Chapter 3

CONFIDENCE BANDS FOR THE RATIO AND ODDS RATIO OF SURVIVAL CURVES UNDER THE COX MODEL

In this chapter, we study simultaneous confidence bands for the ratio and odds ratio of two survival functions under the Cox model. In Section 3.1, we introduce the research motivation. In Section 3.2, major inference results for constructing both Equal precision (EP) and Hall-Wellner (HW) type confidence bands are presented for the ratio of survival functions based on the Cox regression model. In Section 3.3, simultaneous confidence bands are constructed for the survival odds ratio. In Section 3.4, simulation studies are conducted to evaluate the performance of the proposed method. Two real applications demonstrating the utility of the proposed technique are given in Section 3.5, using the primary biliary cirrhosis data from the Mayo Clinic, as well as the chronic myelogenous leukemia data from the International Bone Marrow Transplant Registry and German CML Study Group. Conclusions and discussions are presented in Section 3.6. Proofs are summarized in Appendix A.

3.1 Motivation

In biomedical applications, it is usually of primary desires to compare the survival rates of two treatments (Zhang and Klein [86]). When diagnostic results imply possible associations between some time-varying effects and a certain designated treatment, it is clinically reasonable to consider what times the two treatments differ from each other. Zhang and Klein [86] points out that the question is particularly essential when one treatment has a higher early survival rate but fails in the long term. The authors have also provided a typical example in comparing the survival rates of bone marrow transplantation (BMT) and the traditional chemotherapy patients, where BMT patients might have a higher early mortality rate, but a lower death rate as time goes by. The desire motivates us for using the Cox model in order to incorporate more information from time-varying effects that might eventually affect the survival rates.

However, to evaluate the comparison of two treatment effects in terms of hazard ratios by the Cox model, the absence of proportionality can be problematic. Zhang and Klein [86] has suggested an informal but more straightforward test by the use of graphical methods. Possible choice will include plotting the estimated cumulative (log) baseline hazard rates from stratified Cox models. Andersen [3] develops plots of the estimated cumulative hazard rate from one treatment over another, and Arjas [6] constructed relevant plots based on a modified total time on test statistic.

Several methods have been proposed in literatures for comparing two treatments using different types of functions in the field of survival analysis; for example, Kalbfleisch and Prentice [36], Schemper [69] and Xu and O′Quigley [78] all considered cumulative hazard ratios. Zhang and Klein [86] measured the difference of two survival functions based on the proportional hazards model, while Wei and Schaubel [76] proposed an estimator of the ratio of baseline cumulative hazards in two populations under a stratified Cox model.

To the best knowledge of us, however, methodologies concentrating on simultaneous confidence bands for the ratio or odds ratio of survival functions, in the presence of regression covariates that enable possible adjustments of survival functions, still remain unavailable. In Chapter 3, we propose our solution under the stratified Cox regression model. That is, we estimate the ratio of two survival functions,

$$R(\cdot; z_0) = S_1(\cdot; z_0)/S_2(\cdot; z_0),$$

by

$$\hat{R}(\cdot; z_0) = \hat{S}_1(\cdot; z_0)/\hat{S}_2(\cdot; z_0) = e^{(\hat{\Lambda}_2(\cdot; z_0) - \hat{\Lambda}_1(\cdot; z_0))},$$

and the survival odds ratio

$$OR(\cdot; z_0) = \frac{S_1(\cdot; z_0)}{1 - S_1(\cdot; z_0)} \Big/ \frac{S_2(\cdot; z_0)}{1 - S_2(\cdot; z_0)},$$

by

$$\widehat{OR}(\cdot; z_0) = \frac{\hat{S}_1(t; z_0)}{1 - \hat{S}_1(t; z_0)} \Big/ \frac{\hat{S}_2(t; z_0)}{1 - \hat{S}_2(t; z_0)} = \frac{e^{\hat{\Lambda}_2(t; z_0)} - 1}{e^{\hat{\Lambda}_1(t; z_0)} - 1},$$

under a particular set of covariate values, where $\hat{\Lambda}_i(\cdot; z_0)$, $i = 1,\ 2$ are the Breslow (Breslow [10]) estimators of cumulative hazard functions.

To find simultaneous confidence bands for $R(\cdot; z_0)$ and $OR(\cdot; z_0)$, using the martingale central limit theorem, it can be shown that both

$$W_R(\cdot; z_0) = \sqrt{n}\left[\hat{R}(\cdot; z_0) - R(\cdot; z_0)\right]$$

and

$$W_{OR}(\cdot; z_0) = \sqrt{n}\left[\widehat{OR}(\cdot; z_0) - OR(\cdot; z_0)\right]$$

converge weakly to zero-mean Gaussian processes. It is well-known that the independent increment assumption is not kept by those Gaussian processes. In this chapter, we follow Lin et al. [47] to conduct simulations on our proposed bands using a similar approach.

3.2 Confidence Bands for the Ratio of Two Survival Functions

For subject j in group i, denote T_{ij} as survival time and C_{ij} as censoring time, and assume that T_{ij} and C_{ij} are independent conditional on $Z_{ij}(\cdot)$, where $Z_{ij}(\cdot)$ are bounded covariates. For patients of group i, we fit a Cox regression model (Cox [17]) stratified on treatments specifying the hazard function $\lambda_i(t; z)$ for the failure time T_i under covariate $Z(t) = z(t)$ by the following form

$$\lambda_i(t; z) = \lambda_{i0}(t)e^{\beta' z(t)},$$

where β is a p-vector of unknown regression coefficients, $z(t)$ is a p-vector of possibly time-varying covariates and $\lambda_{i0}(t)$ is the unspecified baseline hazard function.
Therefore, the cumulative hazard function $\Lambda_i(t; z)$ becomes

$$\Lambda_i(t; z) = \int_0^t e^{\beta' z(u)}\lambda_{i0}(u)du.$$

By Breslow [10], for group i, the baseline cumulative hazard function $\Lambda_i(t)$ can be consistently estimated by

$$\hat{\Lambda}_{i0}(t) = \sum_{j=1}^{n_i} \frac{I(X_{ij} \le t)\Delta_{ij}}{\sum_{j=1}^{n_i} Y_{ij}(X_{ij})e^{\hat{\beta}' Z_{ij}(X_{ij})}} \tag{3.1}$$

where $\hat{\beta}$ is the MLE of β and $X_{ij} = \min(T_{ij}, C_{ij})$ is the censored survival time for subject j of group i with $\Delta_{ij} = I(T_{ij} \leq C_{ij})$ being the corresponding censoring indicator. In addition, write $Y_{ij}(t) = I(X_{ij} \geq t)$, indicating if subject j of group i is at risk prior to time t.

Following Lin et al. [47], for group i, we introduce the following notations.

$$\begin{aligned} S_i^r(\beta, t) =& n_i^{-1} \sum_{j=1}^{n_i} Y_{ij}(t) e^{\beta' Z_{ij}(t)} Z_{ij}^{\otimes r}(t), \quad s_i^{(r)} = E\left[S_i^r(\beta, t)\right], \quad r = 0,\ 1,\ 2 \\ E_i(\beta, t) &= \frac{S_i^{(1)}(\beta, t)}{S_i^{(0)}(\beta, t)}, \quad e_i(\beta, t) = \frac{s_i^{(1)}(\beta, t)}{s_i^{(0)}(\beta, t)}, \\ V_i(\beta, t) &= \frac{S_i^{(2)}(\beta, t)}{S_i^{(0)}(\beta, t)} - E_i(\beta, t)^{\otimes 2}, \quad v_i(\beta, t) = \frac{s_i^{(2)}(\beta, t)}{s_i^{(0)}(\beta, t)} - e_i(\beta, t)^{\otimes 2}, \end{aligned}$$

where for a column vector a, $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$ and $a^{\otimes 2} = aa'$.

Define the counting process $N_{ij}(t) = \Delta_{ij} I(X_{ij} \leq t)$ and martingale

$$M_{ij}(t) = N_{ij}(t) - \int_0^t Y_{ij}(u) e^{\beta' Z_{ij}(u)} d\Lambda_{i0}(u), \quad j = 1, \ldots, n_i.$$

In addition, write $\bar{N}_i(u) = \sum_{i=1}^{n_i} N_{ij}(u)$, $\bar{M}_i(u) = \sum_{j=1}^{n_i} M_{ij}(u)$.

Thus, writing in counting processes, (3.1) becomes

$$\hat{\Lambda}_{i0}(t; z_0) = \int_0^t \frac{d\bar{N}_i(u)}{n_i S_i^{(0)}(\hat{\beta}, u).}$$

Theorem 2. *The process*

$$W(t; z_0) = n^{1/2}[(\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)) - (\Lambda_2(t; z_0) - \Lambda_1(t; z_0))]$$

is asymptotically equivalent to

$$\begin{aligned} \widetilde{W}(t; z_0) =& \frac{1}{\sqrt{n}} \frac{1}{p_2} \int_0^t \frac{e^{\beta_0' z_0(u)} d\bar{M}_2(u)}{S_2^{(0)}(\beta_0, u)} - \frac{1}{\sqrt{n}} \frac{1}{p_1} \int_0^t \frac{e^{\beta_0' z_0(u)} d\bar{M}_1(u)}{S_1^{(0)}(\beta_0, u)} \\ &+ (h_2(t; z_0) - h_1(t; z_0))' \Sigma^{-1} \\ &\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \int_0^{\infty} \{Z_{ij}(u) - E_i(\beta_0, u)\} dM_{ij}(u) \right] \end{aligned} \tag{3.2}$$

where $p_i = \lim_{n\to\infty} n_i/n$, *and*

$$h_i(t; z_0) = \int_0^t e^{\beta' z_0}[z_0 - e_i(\beta, u)]d\Lambda_{i0}(u), \qquad i = 1, 2.$$

In addition, its covariance function is given by

$$\begin{aligned}\xi(t, v; z_0) =& \sum_{i=1}^{2} \frac{1}{p_i} \int_0^{t\wedge v} \frac{e^{2\beta_0' z_0(u)} d\bar{N}_i(u)}{n_i \left[s_i^{(0)}(\beta_0, u)\right]^2} \\ &+ (h_1(t; z_0) - h_2(t; z_0))' \Sigma^{-1} (h_1(v; z_0) - h_2(v; z_0)). \end{aligned} \tag{3.3}$$

Since (3.2) does not have independent increment, we follow the method proposed by Lin et al. [47] in order to conduct simulation studies. By replacing $dM_{ij}(t)$ by $G_{ij}dN_{ij}(t)$, where G_{ij}'s are all standard normal random variables, we know $\widetilde{W}(t; z_0)$ can be asymptotically estimated by

$$\begin{aligned}\widehat{W}(t; z_0) =& \frac{1}{\sqrt{n}} \frac{1}{\hat{p}_2} \sum_{j=1}^{n_2} \frac{I(X_{2j} \le t)\Delta_{2j} e^{\hat{\beta} z_0(X_{2j})} G_{2j}}{S_2^{(0)}(\hat{\beta}, X_{1j})} \\ &- \frac{1}{\sqrt{n}} \frac{1}{\hat{p}_1} \sum_{j=1}^{n_1} \frac{I(X_{1j} \le t)\Delta_{1j} e^{\hat{\beta} z_0(X_{1j})} G_{1j}}{S_1^{(0)}(\hat{\beta}, X_{1j})} \\ &+ (\hat{h}_2(t; z_0) - \hat{h}_1(t; z_0))' \widehat{\Sigma}^{-1} \\ &\left[\frac{1}{\sqrt{n}} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \Delta_{ij} \left\{Z_{ij}(X_{ij}) - E_i(\hat{\beta}, X_{ij})\right\} G_{ij}\right] \end{aligned} \tag{3.4}$$

with a consistent estimator of the covariance function written as

$$\begin{aligned}\hat{\sigma}^2(t; z_0) =& \sum_{i=1}^{2} \frac{1}{p_i} \int_0^{t\wedge v} \frac{e^{2\beta_0' z_0(u)} d\bar{N}_i(u)}{n_i \left[S_i^{(0)}(\beta_0, u)\right]^2} \\ &+ (\hat{h}_1(t; z_0) - \hat{h}_2(t; z_0))' \widehat{\Sigma}^{-1} (\hat{h}_1(v; z_0) - \hat{h}_2(v; z_0)), \end{aligned}$$

where,

$$\widehat{\Sigma} = n^{-1} \sum_{i=1}^{2} \sum_{j=1}^{n_i} \Delta_{ij} \{S_i^{(2)}(\hat{\beta}, X_{ij})/S_i^{(0)}(\hat{\beta}, X_{ij}) - E_i(\hat{\beta}, X_{ij})^{\otimes 2}\}, \tag{3.5}$$

and for $i = 1, 2$,

$$\hat{h}_i(t; z_0) = n_i^{-1} \sum_{j=1}^{n_i} I(X_{ij} \leq t) \Delta_{ij} e^{\hat{\beta}' z_0(X_{ij})} \{z_0(X_{ij}) - E_i(\hat{\beta}, X_{ij})\} / S_i^{(0)}(\hat{\beta}, X_{ij}). \tag{3.6}$$

To construct simultaneous confidence bands, we consider the class of transformed process

$$W_R(t; z_0) = n^{\frac{1}{2}} g(t; z_0) \left[\phi\{\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)\} - \phi\{\Lambda_2(t; z_0) - \Lambda_1(t; z_0)\} \right],$$

where ϕ and g maintain the same properties as stated in Lin et al. [47]. Thus, by the functional delta-method, the process $W_R(t; z_0)$ is equivalent to

$$\widetilde{W}_R(t; z_0) = g(t; z_0) \phi'\{\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)\} W(t; z_0),$$

where $\phi'(x)$ is the first derivative of $\phi(x)$.
It is easy to see that the distribution of $\widetilde{W}_R(t; z_0)$ can be approximated by that of

$$\widehat{W}_R(t; z_0) = g(t; z_0) \phi'\{\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)\} \widehat{W}(t; z_0).$$

Let C_α be the upper α-quantile of the distribution $\sup_{t \in [t_1, t_2]} |\widehat{W}_R(t; z_0)|$, which can be estimated through simulation. Then an approximate $100(1-\alpha)\%$ confidence band for $\phi\{\Lambda_2(t; z_0) - \Lambda_1(t; z_0)\}$ over time interval $[t_1, t_2]$ becomes

$$\phi\{\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)\} \pm n^{-\frac{1}{2}} C_\alpha / g(t; z_0).$$

Let $\phi(x) = e^x$. To choose appropriate weight function $g(\cdot; z_0)$, we consider

$$g_1(t; z_0) = \{\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)\} / \hat{\sigma}(t; z_0), \quad g_2(t; z_0) = \{\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)\} / \{1 + \hat{\sigma}^2(t; z_0)\}.$$

The asymptotic $100(1-\alpha)\%$ confidence bands for $R(t;z_0)$ over the time interval $[t_1,t_2]$ are consequently given in the following form,

$$\widehat{R}(t;z_0) \pm n^{-\frac{1}{2}} C_{1,\alpha} \widehat{R}(t;z_0)\hat{\sigma}(t;z_0), \tag{3.7}$$

$$\widehat{R}(t;z_0) \pm n^{-\frac{1}{2}} C_{2,\alpha} \widehat{R}(t;z_0)[1+\hat{\sigma}^2(t;z_0)], \tag{3.8}$$

respectively, where $C_{1,\alpha}$ and $C_{2,\alpha}$ are upper α-quantiles that correspond to g_1 and g_2.
Note that (3.7) and (3.8) are the so-called equal-precision (Nair [58]) and Hall-Wellner (Hall and Wellner [31]) type bands, respectively.

3.3 Confidence Band for the Odds Ratio of Two Survival Functions

Although quite analogous mathematically, since $OR(t;z_0)$ cannot be rewritten into a function of $\{\Lambda_2(t;z_0)-\Lambda_1(t;z_0)\}$, we only give the following theorem. The proof is attached in Appendix A.

Theorem 3. *The process*

$$W_{OR}(t) = n^{1/2}\left[\frac{\hat{S}_1(t;z_0)}{1-\hat{S}_1(t;z_0)}\Big/\frac{\hat{S}_2(t;z_0)}{1-\hat{S}_2(t;z_0)} - \frac{S_1(t;z_0)}{1-S_1(t;z_0)}\Big/\frac{S_2(t;z_0)}{1-S_2(t;z_0)}\right]$$

is asymptotically equivalent to

$$\begin{aligned}\widetilde{W}_{OR}(t;z_0) = {} & \left[\frac{S_1(t;z_0)}{1-S_1(t;z_0)}\Big/\frac{S_2(t;z_0)}{1-S_2(t;z_0)}\right]\cdot \\ & \left\{\frac{1}{\sqrt{n}}\frac{1}{p_2}\frac{1}{1-S_2(t;z_0)}\int_0^t \frac{e^{\beta_0' z_0(u)}d\bar{M}_2(u)}{S_2^{(0)}(\beta_0,u)}\right. \\ & -\frac{1}{\sqrt{n}}\frac{1}{p_1}\frac{1}{1-S_1(t;z_0)}\int_0^t \frac{e^{\beta_0' z_0(u)}d\bar{M}_1(u)}{S_1^{(0)}(\beta_0,u)} \\ & +\left(\frac{h_2(t;z_0)}{1-S_2(t;z_0)}-\frac{h_1(t;z_0)}{1-S_1(t;z_0)}\right)'\Sigma^{-1} \\ & \left.\left[\frac{1}{\sqrt{n}}\sum_{i=i}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(u)-E_i(\beta_0,u)\}dM_{ij}(u)\right]\right\},\end{aligned} \tag{3.9}$$

with the covariate function given by

$$\xi_{OR}(t;z_0)=\left[\frac{S_1(t;z_0)}{1-S_1(t;z_0)}\Big/\frac{S_2(t;z_0)}{1-S_2(t;z_0)}\cdot\frac{S_1(v;z_0)}{1-S_1(v;z_0)}\Big/\frac{S_2(v;z_0)}{1-S_2(v;z_0)}\right]\cdot \left\{\sum_{i=1}^{2}\frac{1}{p_i}\frac{1}{1-S_i(t;z_0)}\frac{1}{1-S_i(v;z_0)}\int_0^{t\wedge v}\frac{e^{2\beta_0' z_0(u)}d\bar{N}_i(u)}{n_i\left[s_i^{(0)}(\beta_0,u)\right]^2} + \left(\frac{h_1(t;z_0)}{1-S_1(t;z_0)}-\frac{h_2(t;z_0)}{1-S_2(t;z_0)}\right)'\Sigma^{-1} \left(\frac{h_1(v;z_0)}{1-S_1(v;z_0)}-\frac{h_2(v;z_0)}{1-S_2(v;z_0)}\right)\right\}. \tag{3.10}$$

Similarly, based on Lin et al. [47], $\widetilde{W}_{OR}(t;z_0)$ can be asymptotically equivalent to

$$\widehat{W}_{OR}(t;z_0)=\left[\frac{\hat{S}_1(t;z_0)}{1-\hat{S}_1(t;z_0)}\Big/\frac{\hat{S}_2(t;z_0)}{1-\hat{S}_2(t;z_0)}\right]\cdot \left\{\frac{1}{\sqrt{n}}\frac{1}{\hat{p}_2}\frac{1}{1-\hat{S}_2(t;z_0)}\sum_{j=1}^{n_2}\frac{I(X_{2j}\le t)\Delta_{2j}e^{\hat{\beta}z_0(X_{2j})}G_{2j}}{\hat{S}_2^{(0)}(\hat{\beta},X_{2j})} - \frac{1}{\sqrt{n}}\frac{1}{\hat{p}_1}\frac{1}{1-\hat{S}_1(t;z_0)}\sum_{j=1}^{n_1}\frac{I(X_{1j}\le t)\Delta_{1j}e^{\hat{\beta}z_0(X_{1j})}G_{1j}}{S_1^{(0)}(\hat{\beta},X_{1j})} + \left(\frac{\hat{h}_2(t;z_0)}{1-\hat{S}_2(t;z_0)}-\frac{\hat{h}_1(t;z_0)}{1-\hat{S}_1(t;z_0)}\right)'\widehat{\Sigma}^{-1} \left[\frac{1}{\sqrt{n}}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\Delta_{ij}\left\{Z_{ij}(X_{ij})-E_i(\hat{\beta},X_{ij})\right\}G_{ij}\right], \tag{3.11}$$

with the consistent estimator of $\xi_{OR}(t;z_0)$ given by

$$\hat{\sigma}^2_{OR}(t;z_0)=\left(\frac{(1-\hat{S}_2(t;z_0))\hat{S}_1(t;z_0)}{\hat{S}_2(t;z_0)(1-\hat{S}_1(t;z_0))}\right)^2\cdot \left\{\sum_{i=1}^{2}\frac{1}{p_i}\left(\frac{1}{1-\hat{S}_i(t;z_0)}\right)^2\int_0^{t}\frac{e^{2\beta_0' z_0(u)}d\bar{N}_i(u)}{n_i\left[S_i^{(0)}(\beta_0,u)\right]^2} + \left(\frac{\hat{h}_1(t;z_0)}{1-\hat{S}_1(t;z_0)}-\frac{\hat{h}_2(t;z_0)}{1-\hat{S}_2(t;z_0)}\right)'\widehat{\Sigma}^{-1} \left(\frac{\hat{h}_1(v;z_0)}{1-\hat{S}_1(v;z_0)}-\frac{\hat{h}_2(v;z_0)}{1-\hat{S}_2(v;z_0)}\right)\right\}, \tag{3.12}$$

where $\widehat{\Sigma}^{-1}$ and $\hat{h}_i(t; z_0)$ are defined in (3.5) and (3.6).

Thus, the $100(1-\alpha)\%$ confidence bands on the time interval $[t_1, t_2]$ for $OR(t; z_0)$ can be written as,

$$\widehat{OR}(t; z_0) \pm n^{-\frac{1}{2}} C_\alpha \hat{\sigma}_{OR}(t; z_0), \tag{3.13}$$

where C_α is the upper α-quantile of the distribution $\sup_{t\in[t_1,t_2]} |\widehat{W}_{OR}(t; z_0)/\hat{\sigma}_{OR}(t; z_0)|$, which can be generated through Monte Carlo methods.

3.4 Simulation Study

To compare the coverage accuracy of our proposed technique at a nominal level of 95%, we design the following series of simulation studies. Let T_{ij}, $j = 1, ..., n_i$, $i = 1, 2$ be the event times, generated via the transformation

$$T_{ij} = \{-\log(U_{ij})/[\alpha_j \exp(\beta_0 Z_{ij})]\}^{1/\gamma_j},$$

where $U_{ij} \sim$ Uniform(0, 1), $\beta_0 = 0.3$ and Z_{ij} are generated from standard normal distribution truncated at ± 5. Thus, $\{T_{ij}\}$ follows a Weibull distribution with hazard function

$$\lambda_{ij}(t) = \alpha_i \gamma_i t^{\gamma_i - 1} \exp\{\beta Z_{ij}\}.$$

By choosing different values of γ_i , $i = 1,\ 2$, we can ensure that the baseline hazard functions for the two groups will not be proportional. Let the censoring times $C_{ij} \sim$ Uniform(2, 4.5), then designated censoring rates can be achieved by varying the value of α_i. For sample sizes $n = 50, 100$, the coverage probability is calculated based on 2,000 simulated samples. For each replicate sample we calculated the 95% simultaneous confidence bands based on (3.7), (3.8) and (3.13).

It is worth to mention that Lin et al. [47] has suggested that, by Nair [58] and Bie et al. [8], one shall restrict the EP band to time interval $[t_1, t_2]$ such that $\hat{c}_1 = 1 - \hat{c}_2 = 0.05$, or 0.1, where

$$\hat{c}_k = \hat{\sigma}^2(t_k; z_0)/\{1 + \hat{\sigma}^2(t_k; z_0)\}.$$

In our case, to estimate two survival functions simultaneously, we have further adjust $\hat{c}_1$ to a moderately larger scale, in order to accommodate both groups. In our simulation studies, we have set $\hat{c}_1 = 0.1$ when $CR = 25\%$, $\hat{c}_1 = 0.2$ when $CR = 50\%$ and $\hat{c}_1 = 0.3$ when $CR = 75\%$. The consequential results are given in Table 3 for the ratio of survival functions and Table 4 for the odds ratio of survival functions. We can see in Table 3 that coverage probabilities of the proposed bands, after appropriate restriction on time interval $[t_1, t_2]$, becomes quite close to the chosen nominal level. For the ratio of two survival functions, HW bands in general tend to be having a bit higher coverage probabilities than those of EP bands. Such finding is further illustrated by real data applications in the next section.

3.5 Real Applications

3.5.1 The Primary Biliary Cirrhosis Data

The Mayo Clinic developed a database for patients with primary biliary cirrhosis (PBC), a fatal chronic live disease. The data is tabulated in the Appendix D.1 of Fleming and Harrington [25]. A total of $n = 312$ patients participated in the randomized clinical trial, where $n_1 = 158$ patients received the treatment (D-penicillamine) and $n_2 = 154$ were treated with a placebo. Censoring (187 of 312) is heavy in the data.

Following Lin et al. [47], we use the same variable transformations for illustration purpose only, and corresponding parameter estimates are provide in their chapter. Moreover, we chose z_0 as the average level of covariate effects, namely, 51 year old, 3.4 gm/dl serum albumin, 1.8 mg/pl serum bilirubin, 10.74 seconds of prothrombin time and no oedema. Figure 1 depicts the Equal Precision and the Hall-Wallner confidence bands for the survival ratio between the case group and the control group. The estimated ratio of survival functions based on Kaplan-Meier curves is also displayed. Figure 2 depicts those of the survival oddsr ratio.

As discussed in the previous section, one has to adjust $\hat{c}_1$ for a more informative output. In our analysis for the PBC data, since the censoring rate rather high (almost 60%), we have set $\hat{c}_1 = 0.2$ for survival ratio and $\hat{c}_1 = 0.23$ for survival odds ratio. Note that on the basis of 10,000 realizations, for survival odds ratio, C_α is found to be 3.17 for the EP band and 0.59 for the HW band; whereas for survival ratio, such critical value becomes 3.02.

Additionally, it is easy to see that HW bands are generally slightly wider than EP bands, which might be resulted from the mathematical properties of corresponding weight functions. Since both simultaneous bands contain the horizontal line marked for identical survival rate (reflected by both ratio and odds ratio in the same sense), no distinctive evidence would imply any difference between the treatment and the placebo on the basis of our analysis.

3.5.2 The Chronic Myelogenous Leukemia Data

The chronic myelogenous leukemia (CML) data consists of patients receiving conventional chemotherapy as well as patients treated by allogeneic bone marrow transplantation. Patients receiving the conventional treatment were from a multicenter trial conducted by the German CML study. Amongst the 196 patients selected within the cohort, 75 received primary treatment with interferon and 121 with hydroxyurea. Patients within this cohort are followed until death or the end of the study.

The transplant cohort contains 548 patients receiving hydroxyurea or interferon pretreatment and a HLA-identical sibling bone marrow transplant (BMT) (Zhang and Klein [86]). The cohort study pertains to the International Bone Marrow Transplant Registry (IBMTR). The IBMTR is a voluntary collaborating group of over 300 transplant centers worldwide that contribute data on their allogeneic bone marrow transplants to a statistical center at the Medical College of Wisconsin. Patients in this group were diagnosed between 1983 and 1991, under ages from 15 to 55. For more details about the cohort study, refer to Gale et al. [26].

Following Zhang and Klein [86], we use the screened cohort consisted by 101 patient with conventional treatment and 399 receiving BMT, whose year of diagnosis is no later than 1988. As pointed out by Zhang and Klein [86], the covariate, namely, spleen size, has heavy joint effect with treatment. Thus, in order to account for such association, the original data is split into a cohort for patients with large spleen size ($\geq$ 10 cm) and that for patients with small spleen size (otherwise). Therefore desired comparisons via survival functions have to be completed separately for the two new cohorts accordingly. Relevant graphical results are shown in Figure 3 to Figure 6.

Carefulness has to be imposed when choosing time interval $[t_1, t_2]$ for the data set. Zhang and Klein [86] has suggested $[6.4, 72.2]$ for survival difference, we adopted the same interval for our

proposed method for survival ratio, but used another interval $[7.9, 72.2]$ to ensure both groups have death occurred.

It can be seen from Figure 3 and Figure 4 that both confidence bands cover an identical line until after about 58 months, indicating that the conventional chemotherapy treatment has an early survival advantage, possibly resulted from toxicity of the BMT. However, the BMT treatment shows a long term survival advantage due to lower relapse rate. It is, nonetheless, interesting to infer from Figure 5 and Figure 6 that the survival odds ratio amongst the two treatments is quite likely to be relatively a constant. Such proportionality hypothesis has also been suggested by a form test conducted by Zhang and Klein [86], and yet a graphical illustration might seem to be more revealing.

3.6 Discussion

It is known that the simultaneous confidence band for survival curves is more appealing than pointwise confidence interval, in the sense that it measures and demonstrates the overall trend of survival rates over a period of time, which will be more likely to reveal a more thorough and comprehensive clinical reference. For analysis involving the proportional odds regression model, which is of a great capacity of applications (Yang and Prentice [78, 79]), it is analytically worthwhile to build visual illustration to form a graphical test, and hence constructing confidence bands for the survival odds ratio becomes statistically critical.

In addition, The estimated critical value, C_α, depends on the number of realizations, N. Thus, it is critical to know an appropriate N for applying our proposed method. Parzen et al. [62] has reported a cut-off value of N for constructing simultaneous confidence interval for the difference of two survival functions, while Zhang and Klein [86] has suggested that $N = 500$ for their simultaneous confidence band construction for survival difference. In our setting, we need at least 500 iterations to obtain stable critical values for the survival ratio, however, N has to be at least 1000 for the survival odds ratio.

Last but not least, it is well-known that in many survival literatures, the Box-Cox transformation is commonly used for the cumulative hazard function in order to obtain more realistic survival estimate, which will more likely result in more statistically reasonable results. Note that such transformation becomes mathematically improbable for the odds ratio of survival functions. Similar to

Lin et al. [47], by letting $\phi(x) = \log(x)$, one can easily derive the log-transformed confidence bands for the ratio of survival functions. Our simulation studies, nevertheless, show quite similar results between transformed and untransformed bands. Consequently, for contextual integrity, we decide to use untransformed bands consistently.

Chapter 4

CONFIDENCE BANDS FOR THE RATIO AND ODDS RATIO OF SURVIVAL CURVES UNDER THE ADDITIVE RISK MODEL

In this chapter, we extend the methodologies described in Chapter 3 toward the additive risk model to construct simultaneous confidence bands for $R(\cdot; z_0)$ and $OR(\cdot; z_0)$. Similar to what is discussed in Chapter 3, $W_R(\cdot; z_0)$ and $W_{OR}(\cdot; z_0)$ under the additive risk model still do not have the independent increment structure, and therefore cannot be transformed into standard Brownian bridges. Based on Lin and Ying [51], we show an analogous method to construct the simultaneous confidence bands .

In the Section 4.2, we provide notations and a review of the semiparametric additive risk model. In Section 4.3, we provide a simulation study for constructing the simultaneous confidence bands for the ratio and odds ratio of two survival curves. Two clinical trial data sets are studied in Section 4.4 for further illustration. Some discussions are given in Section 4.5.

4.1 Literature Review

For two sample comparison incorporating time-variant adjustments, one possible choice would be the Cox proportional hazards model (Cox [17]), which specifies the association between the hazard function and the covariates with an exponential link function. More specifically, the Cox model specifies the hazard rate to be of the following form,

$$\lambda(t; \mathbf{Z}_i) = \lambda_0(t) \exp\{\beta' \mathbf{Z}_i(t)\},$$

where $\mathbf{Z}_i(t)$ is the ith possible time-dependent unknown covariate vector and β is the regression parameter. Because of the multiplicative association between covariate effects and the baseline hazard function, the infinite dimensional nuisance parameter $\lambda_0(t)$ cancels out in the partial likelihood structure (Cox [18]). In some applications, when the multiplicative effect assumption of regression covariates on the hazard function is violated, the additive risk model provides a meaningful alter-

native. The additive risk model, proposed by Lin and Ying [51], assumes linearity between the covariate and the hazard function,

$$\lambda(t; \mathbf{Z}_i) = \lambda_0(t) + \beta' \mathbf{Z}_i(t).$$

As stated in Chapter 1, both the Cox model and the additive risk models have extensive biological applications and rigorous statistical foundations. Jointly, the two models provide equally important reference to modeling the hazard function (Breslow and Day [11, 12]). Lin and Ying [51] provided a semiparametric solution to the additive risk model with closed form estimation of regression coefficients, and further derived large-sample properties for this model. Since the additive risk model might in some cases be more appropriate than the Cox model, it plays a key role as an important modeling alternative, although model checking and diagnosis theories are not yet completed. Recent literatures have shown an increasing popularity of applications through the additive risk model toward a variety of problems, on the basis of counting processes and martingale. For instance, Lin and Ying [52] and Lin et al. [48] analyzed the interval censored data, Shen and Cheng [72] studied the cumulative incidence curve in the context of competing risks, Yip et al. [83] designed recapture experiments and Kulich and Lin [40] proposed methodologies in evaluating measurement errors. Song et al. [74] and Yin and Hu [81] both studied confidence bands of survival functions under the additive risk model. Quite recently, Lee and Hyun [42] have proposed confidence bands for the difference of two survival functions.

Many other researches have also been conducted relevant to the original additive risk model of Lin and Ying [51], which, among others, include Aalen [1, 2] and Huffer and McKeague [34] for nonparametric additive risk models, McKeague and Sasieni [54] for a partly additive risk model and Scheike [68] for the rate function based on both nonparametric and semiparametric additive risks. These models allow adjustments for time-varying effects which provide more flexibility in model fitting.

The estimation of parameters based on the additive risk model has an analytic closed form that is computationally easy to implement. The risk ratio and the risk odds ratio are alternatives to the risk difference to assess comparison between treatments in epidemiological studies. The ratio

and odds ratio of survival functions carry out comparisons on the chance of death pertaining to two sets of characteristics.

Compared to confidence interval, confidence band is more informative and hence has become more desirable for describing the entire survival experience. In the nonparametric setup without considering any adjustable covariate, the confidence bands, namely the equal precision band and Hall-Wellner type band, have been studied extensively and are described in great detail by Fleming and Harrington [25] and Andersen et al. [5].

4.2 Construction of Confidence Bands for Comparing Two Survival Functions

Adopting the same settings as those in Chapter 3, for a patient j in group i, $i = 1, 2$, we fit an additive risk model (Lin and Ying [51]) stratified on treatment. The hazard function has the following form

$$\lambda_i(t; Z_{ij}) = \lambda_{i0}(t) + \beta' Z_{ij}(t), \tag{4.1}$$

where $\lambda_{i0}(t)$ is the unspecified baseline hazard function for group i, $Z_{ij}(\cdot)$ is a p-vector of covariates that influence the hazard rate and hence the survival rate, and β is a p-vector of regression coefficients.

For group i with n_i independent subjects, we consider the counting process $\{N_{ij}(t); t \geq 0\}$ for the jth subject in the group that records the number of observed events up to time t. The intensity function for $N_{ij}(t)$ is thus given by

$$Y_{ij}(t)\ d\Lambda(t; Z_{ij}) = Y_{ij}(t)\{d\Lambda_{i0}(t) + \beta' Z_{ij}(t)dt\}, \tag{4.2}$$

where $Y_{ij}(t) = I(X_{ij} \geq t)$, and

$$\Lambda_{i0}(t) = \int_0^t \lambda_{i0}(u)du.$$

The counting process $N_{ij}(\cdot)$ can be uniquely decomposed such that for every j and t,

$$N_{ij}(t) = M_{ij}(t) + \int_0^t Y_{ij}(u)\ d\Lambda(u; Z_{ij}(u)), \tag{4.3}$$

where $M_{ij}(\cdot)$ is a local square integrable martingale (Lin and Ying [51]).

Estimators for baseline cumulative hazard functions are given by

$$\hat{\Lambda}_{i0}(t;\hat{\beta}) = \int_0^t \frac{\sum_{j=1}^{n_i}\{dN_{ij}(u) - Y_{ij}(u)\hat{\beta}' Z_{ij}(u)du\}}{\sum_{j=1}^{n_i} Y_{ij}(u)}. \tag{4.4}$$

Given a covariate vector $z_0(t)$, the survival function can be estimated by

$$\hat{S}_i(t;z_0) = \exp\left\{-\hat{\Lambda}_{i0}(t,\hat{\beta}) - \hat{\beta}\int_0^t z_0(u)du\right\}, \tag{4.5}$$

In addition, β can be estimated from the following estimating equation,

$$U(\beta) = \sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty Z_{ij}(t)\{dN_{ij}(t) - Y_{ij}(t)d\hat{\Lambda}_{i0}(\beta,t) - Y_{ij}(t)\beta' Z_{ij}(t)dt\},$$

which is equivalent to

$$U(\beta) = \sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(t) - \bar{Z}_i(t)\}\{dN_{ij}(t) - Y_{ij}(t)\beta' Z_{ij}(t)dt\}, \tag{4.6}$$

where

$$\bar{Z}_i(t) = \sum_{j=1}^{n_i}\frac{Y_{ij}(t)}{\sum_{j=1}^{n_i} Y_{ij}(t)} Z_{ij}(t).$$

The resulting estimator takes the explicit form

$$\hat{\beta} = \left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty Y_{ij}(t)\{Z_{ij}(t) - \bar{Z}_i(t)\}^{\otimes 2}dt\right]^{-1} \left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(t) - \bar{Z}_i(t)\}dN_{ij}(t)\right]. \tag{4.7}$$

For notational simplicity and convenience, we write

$$\bar{Y}_i(t) = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}(t), \quad i = 1, 2$$
$$G_i(t; z_0) = \int_0^t \{z_0(s) - \bar{Z}_i(s)\}ds, \quad i = 1, 2$$
$$\Sigma = \left[\frac{1}{n}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(t) - \bar{Z}_i(t)\}^{\otimes 2} Y_{ij}(t)dt\right]'.$$

To construct confidences band for the survival ratio $R(t; z_0)$ and odds ratio $OR(t; z_0)$, we introduce the following theorems. Proofs are given in Appendix A.

Theorem 4. *Under the additive risk model, for the survival ratio, the process*

$$L(t; z_0) = n^{1/2}[(\hat{\Lambda}_2(t; z_0) - \hat{\Lambda}_1(t; z_0)) - (\Lambda_2(t; z_0) - \Lambda_1(t; z_0))]$$

converges weakly to the Gaussian process

$$\begin{aligned}\widetilde{L}(t; z_0) =& \frac{1}{\sqrt{n}}\frac{1}{p_2}\sum_{j=1}^{n_2}\int_0^t \frac{1}{\bar{Y}_2(s)}dM_{2j}(s) - \frac{1}{\sqrt{n}}\frac{1}{p_1}\sum_{j=1}^{n_1}\int_0^t \frac{1}{\bar{Y}_1(s)}dM_{1j}(s) \\ &+ \frac{G'(t; z_0)}{\sqrt{n}}\Sigma^{-1}\left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(s) - \bar{Z}_i(s)\}dM_{ij}(s)\right], \end{aligned} \tag{4.8}$$

with the covariance function estimated by

$$\begin{aligned}\xi_L(t; z_0) = \sum_{i=1}^{2}\Bigg[&\frac{1}{n}\frac{1}{p_i^2}\sum_{j=1}^{n_i}\int_0^{t\wedge s}\frac{dN_{ij}(u)}{(\bar{Y}_i(u))^2} + (-1)^i\frac{G'(s; z_0)}{p_i}\Sigma^{-1}D_{i1}(t) \\ &+ (-1)^i\frac{G'(t; z_0)}{p_i}\Sigma^{-1}D_{i1}(s)\Bigg] + G'(t; z_0)\Sigma^{-1}D_{i2}(\Sigma^{-1})'G(s; z_0),\end{aligned}$$

where

$$G(t; z_0) = G_2(t; z_0) - G_1(t; z_0), \tag{4.9}$$

$$D_{i1}(t) = \sum_{j=1}^{n_i}\int_0^t \frac{\{Z_{ij}(u) - \bar{Z}_i(u)\}dN_{ij}(u)}{\sum_{k=1}^{n_i} Y_{ik}(u)}, \tag{4.10}$$

$$D_{i2} = \frac{1}{n}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(t) - \bar{Z}_i(t)\}^{\otimes 2} dN_{ij}(t), \tag{4.11}$$

and for column vector a, $a^{\otimes 2}$ denotes the outer product of a.

Theorem 5. *The process of the survival odds ratio,*

$$L_{OR}(t) = n^{1/2}\left[\frac{\hat{S}_1(t;z_0)}{1-\hat{S}_1(t;z_0)}\Big/\frac{\hat{S}_2(t;z_0)}{1-\hat{S}_2(t;z_0)} - \frac{S_1(t;z_0)}{1-S_1(t;z_0)}\Big/\frac{S_2(t;z_0)}{1-S_2(t;z_0)}\right]$$

is asymptotically equivalent to

$$\begin{aligned}\widetilde{L}_{OR}(t;z_0) =& \frac{(1-S_2(t;z_0))S_1(t;z_0)}{S_2(t;z_0)(1-S_1(t;z_0))}\cdot \\ & \left\{\frac{1}{\sqrt{n}}\frac{1}{p_2(1-S_2(t;z_0))}\sum_{j=1}^{n_2}\int_0^t \frac{1}{\bar{Y}_2(s)}dM_{2j}(s)\right. \\ & -\frac{1}{\sqrt{n}}\frac{1}{p_1(1-S_1(t;z_0))}\sum_{j=1}^{n_1}\int_0^t \frac{1}{\bar{Y}_1(s)}dM_{1j}(s) \\ & +\frac{1}{\sqrt{n}}G'_{OR}(t;z_0)\Sigma^{-1} \\ & \left.\left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(s)-\bar{Z}_i(s)\}dM_{ij}(s)\right]\right\},\end{aligned} \tag{4.12}$$

with covariance function estimated by

$$\begin{aligned}\xi_{L_{OR}}(t;z_0) = & \frac{(1-S_2(t;z_0))S_1(t;z_0)}{S_2(t;z_0)(1-S_1(t;z_0))}\frac{(1-S_2(s;z_0))S_1(s;z_0)}{S_2(s;z_0)(1-S_1(s;z_0))}\cdot \\ & \left\{\sum_{i=1}^{2}\left[\frac{1}{n}\frac{1}{p_i^2}\frac{1}{(1-S_i(t;z_0))}\frac{1}{(1-S_i(s;z_0))}\sum_{j=1}^{n_i}\int_0^{t\wedge s}\frac{dN_{ij}(u)}{(\bar{Y}_i(u))^2}\right.\right. \\ & +(-1)^i\frac{G'_{OR}(s;z_0)}{(1-S_i(t;z_0))p_i}\Sigma^{-1}D_{i1}(t) \\ & \left.+(-1)^i\frac{G'_{OR}(t;z_0)}{p_i(1-S_i(s;z_0))}\Sigma^{-1}D_{i1}(s)\right] \\ & \left.+G'_{OR}(t;z_0)\Sigma^{-1}D_{i2}(\Sigma^{-1})'G_{OR}(s;z_0)\right\},\end{aligned}$$

where

$$G_{OR}(t;z_0) = \frac{G_2(t;z_0)}{1-S_2(t;z_0)} - \frac{G_1(t;z_0)}{1-S_1(t;z_0)}.$$

It is well known that processes $\widetilde{L}(t;z_0)$ and $\widetilde{L}_{OR}(t;z_0)$ do not have independent increments, even if covariates are time invariant, and thus the limiting distributions cannot be transformed to the standard Brownian bridge structure for the construction of simultaneous confidence bands. Lin et al. [47] provides a useful technique by replacing $M_{ij}(u)$ by $N_{ij}(u)G_{ij}$, where $N_{ij}(u)$'s are observed counting processes and G_{ij}'s are i.i.d standard normal random variables. Thereafter, $\widetilde{L}(t;z_0)$ and $\widetilde{L}_{OR}(t;z_0)$ have the same limit distribution with

$$\begin{aligned}\widehat{L}(t;z_0) =& \frac{1}{\sqrt{n}}\frac{1}{\hat{p}_2}\sum_{j=1}^{n_2}\frac{1}{\bar{Y}_2(X_{2j})}I(X_{2j}\le t)\Delta_{2j}G_{2j} \\ &- \frac{1}{\sqrt{n}}\frac{1}{\hat{p}_1}\sum_{j=1}^{n_1}\frac{1}{\bar{Y}_1(X_{1j})}I(X_{1j}\le t)\Delta_{1j}G_{2j} \\ &+ \frac{G'(s;z_0)}{\sqrt{n}}\Sigma^{-1}\left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\{Z_{ij}(X_{ij})-\bar{Z}_i(X_{ij})\}\Delta_{ij}G_{ij}\right].\end{aligned}$$

and with the consistent estimator of the covariance function being

$$\begin{aligned}\hat{\sigma}_L^2(t;z_0) =& \sum_{i=1}^{2}\left[\frac{1}{n}\frac{1}{\hat{p}_i^2}\sum_{j=1}^{n_i}\frac{I(X_{ij}\le t)\Delta_{ij}}{(\bar{Y}_i(X_{ij}))^2} + 2(-1)^i\frac{G'(t;z_0)}{\hat{p}_i}\Sigma^{-1}\widehat{D}_{i1}(t)\right] \\ &+ G'(t;z_0)\Sigma^{-1}\widehat{D}_{i2}(\Sigma^{-1})'G(t;z_0).\end{aligned}$$

And

$$\begin{aligned}\widehat{L}_{OR}(t;z_0) =& \frac{(1-\hat{S}_2(t;z_0))\hat{S}_1(t;z_0)}{\hat{S}_2(t;z_0)(1-\hat{S}_1(t;z_0))} \cdot \\ & \left\{ \frac{1}{\sqrt{n}} \frac{1}{\hat{p}_2(1-\hat{S}_2(t;z_0))} \sum_{j=1}^{n_2} \frac{1}{\bar{Y}_2(X_{1j})} I(X_{2j} \le t)\Delta_{2j} G_{2j} \right. \\ & - \frac{1}{\sqrt{n}} \frac{1}{\hat{p}_1(1-\hat{S}_1(t;z_0))} \sum_{j=1}^{n_1} \frac{1}{\bar{Y}_1(X_{1j})} I(X_{1j} \le t)\Delta_{1j} G_{1j} \\ & + \frac{1}{\sqrt{n}} G'_{OR}(t;z_0)\Sigma^{-1} \\ & \left. \left[\sum_{i=1}^{2} \sum_{j=1}^{n_i} \{Z_{ij}(X_{ij}) - \bar{Z}_i(X_{ij})\}\Delta_{ij} G_{ij} \right] \right\},\end{aligned}$$

with covariance function

$$\begin{aligned}\hat{\sigma}^2_{L_{OR}}(t;z_0) = & \left(\frac{(1-\hat{S}_2(t;z_0))\hat{S}_1(t;z_0)}{\hat{S}_2(t;z_0)(1-\hat{S}_1(t;z_0))} \right)^2 \cdot \\ & \left\{ \sum_{i=1}^{2} \left[\frac{1}{n} \frac{1}{\hat{p}_i^2} \frac{1}{(1-\hat{S}_i(t;z_0))^2} \sum_{j=1}^{n_i} \frac{I(X_{ij} \le t)\Delta_{ij}}{(\bar{Y}_i(X_{ij}))^2} \right. \right. \\ & \left. + 2(-1)^i \frac{G'_{OR}(t;z_0)}{(1-\hat{S}_i(t;z_0))\hat{p}_i} \Sigma^{-1} \widehat{D}_{i1}(t) \right] \\ & \left. + G'_{OR}(t;z_0)\Sigma^{-1} \widehat{D}_{i2} (\Sigma^{-1})' G_{OR}(t;z_0) \right\}.\end{aligned}$$

respectively,

where $\hat{S}_i(t;z_0)$ is given by (4.5), $\hat{p}_i = n_i/n$,

$$\widehat{D}_{i1}(t) = \sum_{j=1}^{n_i} \frac{\{Z_{ij}(X_{ij}) - \bar{Z}_i(X_{ij})\} I(X_{ij} \le t)\Delta_{ij}}{\sum_{k=1}^{n_i} Y_{ik}(X_{ij})},$$

and

$$\widehat{D}_{i2} = \frac{1}{n} \sum_{j=1}^{n_i} \{Z_{ij}(X_{ij}) - \bar{Z}_i(X_{ij})\}^{\otimes 2} \Delta_{ij}.$$

Therefore, the $100(1-\alpha)\%$ EP and HW type confidence bands for $R(t; z_0)$ under the additive risk model over the time interval $[t_1, t_2]$ can be expressed as,

$$\widehat{R}(t; z_0) \pm n^{-\frac{1}{2}} C_{1,\alpha} \widehat{R}(t; z_0) \hat{\sigma}_L(t; z_0),$$
$$\widehat{R}(t; z_0) \pm n^{-\frac{1}{2}} C_{2,\alpha} \widehat{R}(t; z_0)[1 + \hat{\sigma}_L^2(t; z_0)],$$

respectively, where $C_{1,\alpha}$ and $C_{2,\alpha}$ are defined similarly as those in (3.7) and (3.8) of Chapter 3, which can be generated through Monte Carlo methods.

For the survival odds ratio $OR(t; z_0)$, the $100(1-\alpha)\%$ confidence band is given by

$$\widehat{OR}(t; z_0) \pm n^{-\frac{1}{2}} C_\alpha \hat{\sigma}_{L_{OR}}(t; z_0),$$

where C_α is the upper α-quantile of the distribution $\sup_{t \in [t_1, t_2]} |\widehat{L}_{OR}(t; z_0)/\hat{\sigma}_{L_{OR}}(t; z_0)|$, which can be generated through Monte Carlo methods..

4.3 Simulation Study

To evaluate the properties of our proposed method for finite sample sizes, we carry out the following simulation studies. More specifically, Let T_{ij}, $j = 1, ..., n_i$, $j = 1, 2$ be the event times, generated via the transformation

$$T_{ij} = -\log(U_{ij})/(1 + \beta_0 Z_{ij}),$$

where $U_{ij} \sim$ Uniform(0, 1), $\beta = 0.3$ and Z_{ij} are generated from standard normal distribution truncated at ± 5. Thus, $\{T_{ij}\}$ follows an Exponential distribution with hazard function

$$\lambda_{ij}(t) = 1 + \beta_0 Z_{ij}(t).$$

Let the censoring times $C_{ij} \sim$ Uniform(0, 4) to generate 25% censoring rate, Uniform(0, 2) for 50% and Uniform(0, 1) for 75%. For sample sizes $n = 50, 100$, the coverage probability is calculated based on 1,000 simulated samples. For each replicate sample we construct 95% simultaneous confidence

bands for both the survival ratio and odds ratio. The study results are summarized in Table 5 for survival ratio and Table 6 for survival odds ratios.

Note that by Nair [58] and Bie et al. [8], the EP band should be restricted to the time interval $[t_1, t_2]$ such that $\hat{c}_1 = 1 - \hat{c}_2 = 0.05$, or 0.1, where

$$\hat{c}_k = \hat{\sigma}^2(t_k; z_0)/\{1 + \hat{\sigma}^2(t_k; z_0)\}.$$

In our case, to estimate two survival functions simultaneously, we have further adjust $\hat{c}_1$ to a moderately larger scale, in order to account for both groups. More specifically, we have set $\hat{c}_1 = 0.1$ when $CR = 25\%$, $\hat{c}_1 = 0.2$ when $CR = 50\%$ and $\hat{c}_1 = 0.3$ when $CR = 75\%$. The consequential results are given in Table 5 for the ratio of survival functions and Table 6 for the odds ratio of survival functions. We can see in Table 5 that coverage probabilities of the proposed bands, after appropriate restriction on time interval $[t_1, t_2]$, becomes quite close to the chosen nominal level. For the ratio of two survival functions, HW bands in general tend to be having a bit higher coverage probabilities than those of EP bands. Confidence bands for the odds ratio of survival functions are generally wider than those for the ratio. Such finding is further illustrated by real data applications in the next section.

4.4 Real Application

4.4.1 The CML Data

We try the same CML data as that in Chapter 3. The CML data consists of patients receiving conventional chemotherapy as well as patients treated by allogeneic bone marrow transplantation. There are 196 patients receiving the conventional chemotherapy treatment and 548 patients receiving the HLA-identical sibling bone marrow transplant (BMT).

In order to account for the joint effect between treatment and the spleen size, we apply the same analytical strategy as described in Chapter 3. Relevant graphical results are shown in Figure 7 to Figure 10.

It seems, by comparing Figure 3 and Figure 4 with Figure 7 and Figure 8, that the additive risk model might be a more appropriate modeling assumption for the CML data in the sense that the

confidence bands shown are relatively narrower than those in Chapter 3. However, the performance of our proposed method is relatively poor for the survival odds ratio for patients with small spleen size.

4.4.2 The Gastric Cancer Data

The Gastrointestinal Tumor Study Group [27] reported the results of a trial that compared chemotherapy with combined chemotherapy and radiation therapy in the treatment of locally unresectable gastric cancer (Yang and Prentice [80]). There were 45 patients on each treatment group. Censoring rate was relatively low, 4% in the chemotherapy group and 13.33% in the combination group. Estimated survival curves of the two groups intersect at around 33 months after diagnosis. To fit the new model to the data, let the dummy variable Z_i be zero for the chemotherapy group and one for the combination group. The estimated $\hat{\beta} = -0.0012$. We choose $z_0 = 1$. C_α is calculated by 5000 realizations of $\widehat{L}(t; z_0)$. For the survival ratio, the critical value is estimated to be 4.405 for the EP band and 1.843 the HW band, yet for the survival odds ratio, such value jumps to 22.410. Graphical results are given in Figure 11 and Figure 12 for the survival ratio and odds ratio, respectively. It can be seen from Figure 11 that confidence bands are a bit too wide. Such issue appears even worse for the survival odds ratio shown in Figure 12. It is possible that the additive risk model is not fitted here. However, since both plots contain the identical line inside the confidence bands, one might be able to conclude that there is no distinctive difference between the two treatments investigated.

4.5 Remarks

As they stand, (4.4) and (4.5) may not be always monotone in t. We introduce,

$$\hat{\Lambda}^*_{i0}(t) = \max_{s \le t} \hat{\Lambda}_{i0}(\hat{\beta}, s), \quad \hat{S}^*_i(t; z) = \min_{s \le t} \hat{S}(s; z), \quad i = 1, 2.$$

Similar to the argument of Lin and Ying [51], we can show that $\hat{\Lambda}^*_{i0}(t) - \hat{\Lambda}_{i0}(t) = o_p(n_i^{\frac{1}{2}})$. Since $\lim n_i/n = p_i \in (0, 1)$, we know $\hat{\Lambda}^*_{i0}(t) - \hat{\Lambda}_{i0}(t) = o_p(n^{\frac{1}{2}})$, and hence $n^{\frac{1}{2}}(\hat{\Lambda}^*_{i0}(t) - \Lambda_{i0}(t))$ converges to the same limiting distribution as $n^{\frac{1}{2}}(\hat{\Lambda}_{i0}(t) - \Lambda_{i0}(t))$. Then, by substituting $\hat{\Lambda}^*_{i0}(t)$ for $\hat{\Lambda}_{i0}(t)$, we can show that Theorem 4 and Theorem 5 still hold, while the monotonicity is ensured. This modification

is particularly useful when analyzing real clinical trial data sets. For the gastric cancer data, the performance of our proposed method is improved significantly after applying the modification.

Another concern for the proposed model is that $\beta' Z(t)$ should be not too large because otherwise (4.4) will be negative, and (4.5) will exceed one. Lin and Ying [51] suggest a solution by substituting $\exp\{\beta' Z(t)\}$ for $\beta' Z(t)$. We apply the suggestion in several data analysis but discover that in doing so might significantly affect the estimation of the survival rate. In some of the data analysis, the estimated survival rate is 50% less than its that without taking the exponential. We believe the more general solution, by introducing the general regression function $g(\beta' Z(t))$, should be more realistic in most situations due to its modeling flexibility. However, the resulting procedures might not still retain those good properties of the linear form and therefore one might need to use numerical algorithms in solving $\hat{\beta}$ since there might not be a closed form solution to the estimating equation anymore.

Chapter 5

ADJUSTED ATTRIBUTABLE FRACTION FUNCTION FOR CENSORED TIME-TO-EVENT UNDER THE ADDITIVE RISK MODEL

In the present chapter, we study a semiparametric estimation of the population attributable fraction function with the censored time-to-event under the additive risk model. We adjust the semiparametric estimator for the attributable fraction function for practical flexibility to take into account adjusted risk factors that are either discrete or continuous and possibly time-dependent. The introductory methodology is given in what follows, but the proof will be described in the future work.

5.1 Motivation and Literature Review

An important task in public health research is to assess the excess risk attributable to an exposure in a given population (Chen et al. [16]). The population parameter that characterizes the attributable risk is normally regarded as the population attributable fraction. The preposition of the population attributable fraction is first given by Levin [43]. The population attributable fraction, according to Rothman and Greenland [67], is defined as 'the reduction in incidence that would be achieved if the population had been entirely unexposed, compared with its current (actual) exposure pattern'. These measures have received considerable attention in recent years (Benichou [7], Greenland [29], Sliverberg et al. [73], Graubard and Fears [28], Chen et al. [16] and Chen et al. [15]).

Let D be a binary disease status and B be a binary exposure indicator. The population attributable fraction is defined as (Levin [43])

$$A = \frac{\mathrm{P}(D=1) - \mathrm{P}(D=1|B=0)}{\mathrm{P}(D=1)}.$$

This measure is defined for the binary exposure factor only. In the presence of confounding by other risk factors, say, W_k, $k = 1, 2, 3, \ldots, p$, it is more appropriate to use the adjusted attributable

fraction

$$A_{adj} = \frac{\mathrm{P}(D=1) - \sum_{k=1}^{p} \mathrm{P}(W_k = w_k)\mathrm{P}(D=1|B=0,\ W_k = w_k)}{\mathrm{P}(D=1)},$$

where $w_1, \ldots, w_p$ are the corresponding levels of $W_1, \ldots, W_p$ (Bruzzi et al. [13], Whittemore [77]).

The aforementioned measurements are defined for binary outcomes. However, they might not be adequate enough for cohort or clinical trial studies which commonly record censored time-to-event and possibly time-dependent risk factors. Chen et al. [16] first extended the population attributable fraction to a function of the censored time-to-event by replacing the disease incidence rate with the cumulative distribution function of the censored time-to-event. They proposed an estimator for the population attributable fraction function under the Cox model. Chen et al. [15] further established a more comprehensive analysis about both the unadjusted and adjusted population attributive fraction function based on transformation models. As an important modeling alternative, the additive risk model is more appropriate to some clinical trial data set, and features a much easier practical implementation both inferentially and computationally. In order to incorporate time-dependent risk factors, we propose our method for estimating the adjusted attributable fraction function under the additive risk model.

5.2 Inference procedures

For subject i, let T_i and C_i be the time-to-event and the censoring time that are independent conditional on $\left(B_i, W_i^T(\cdot)\right)^T$, where $W_i(\cdot)$ is a p-vector representing possible time-varying covariates and the binary variable B_i is the exposure indicator. Then, suppose the data is recorded in the form of independent $(X_i, \Delta_i, (B_i, W_i^T(t))^T)$, where $X_i = \min(T_i, C_i)$, $\Delta_i = I(T_i \leq C_i)$ and $I(\cdot)$ is the indicator function. We fit an additive risk model (Lin and Ying [51]), in which the hazard function has the following form

$$\lambda(t; B_i, W_i) = \lambda_0(t) + \beta^T \left(B_i, W_i^T(t)\right)^T,$$

where $\lambda_0(t)$ is the unspecified baseline hazard function and β_{p+1} is the vector of regression coefficients.

The adjusted population attributable fraction function is defined as

$$A(t) = \frac{\mathrm{P}(T \leq t) - E\left[\mathrm{P}(T \leq t|B = 0, W = w)\right]}{P(T \leq t)},$$

Chen et al. [15] expresses A(t) in terms of survival functions,

$$A(t) = \frac{S_0(t) - S(t)}{1 - S(t)},$$

where $S(t) = \mathrm{P}(T > t)$ and $S_0(t) = E_W\left[\mathrm{P}(T > t|(0, W^T))\right]$. We show in the sequel that $S_0(t)$ and $S(t)$ are estimated by semiparametric estimators under the additive risk model, $\hat{S}_0(t) = n^{-1}\sum_{i=1}^{n}\hat{S}\{t|(0, W_i^T)\}$ and $\hat{S}(t) = n^{-1}\sum_{i=1}^{n}\hat{S}\{t|(B_i, W_i^T)$, respectively, where $\hat{S}(\cdot|(B_i, W_i^T))$ is a semiparametric estimator of $S(t|(B_i, W_i^T)$ under the additive risk model, and W_i is the observation of subject i for W. Then $A(t)$ is naturally estimated by

$$\hat{A}(t) = \frac{\hat{S}_0(t) - \hat{S}(t)}{1 - \hat{S}(t)}.$$

In order to construct confidence interval and confidence band, it is mathematically convenient to introduce the counting process and martingale framework. Consider for a set of n independent subjects, the counting process $\{N_i(t) = I(X_i \leq t, \Delta_i = 1); t \geq 0\}$ for the ith subject in the group records the number of observed events up to time t. The intensity function for $N_i(t)$ is given by

$$Y_i(t)\ d\Lambda(t; Z_i) = Y_i(t)\{d\Lambda_0(t) + \beta^T\left(B_i, W_i^T(t)\right)^T dt\},$$

where $Y_i(t) = I(X_i \geq t)$, and

$$\Lambda_0(t) = \int_0^t \lambda_0(u)du.$$

The counting process $N_i(\cdot)$ can be uniquely decomposed such that for every i and t,

$$N_i(t) = M_i(t) + \int_0^t Y_i(u)\ d\Lambda\left(u; B_i, W_i^T(u)\right),$$

where $M_i(\cdot)$ is a local square integrable martingale (Lin and Ying [51]).

By Lin and Ying [51], it is easy to see that the estimators for the baseline cumulative hazard functions is given by

$$\hat{\Lambda}_0(t;\hat{\beta}) = \int_0^t \frac{\sum_{i=1}^n \{dN_i(u) - Y_i(u)\beta^T \left(B_i, W_i^T(t)\right)^T du\}}{\sum_{i=1}^n Y_i(u)}. \tag{5.1}$$

Therefore, based on the prequel, we know

$$\hat{S}(t) = n^{-1}\sum_{i=1}^n \exp\left\{-\hat{\Lambda}_0(t,\hat{\beta}) - \hat{\beta}^T \int_0^t \left(B_i, W_i^T(u)\right)^T du\right\}, \tag{5.2}$$

Note that when $E = 0$,

$$\hat{S}_0(t) = n^{-1}\sum_{i=1}^n \exp\left\{-\hat{\Lambda}_0(t,\hat{\beta}) - \hat{\beta}^T \int_0^t \left(0, W_i^T(u)\right)^T du\right\}. \tag{5.3}$$

For notional simplicity, write $Z_i(t) = \left(E_i, W_i^T(t)\right)_{p+1}^T$ as the vector for the exposure indicator and time-varying covariates combined.

β can be estimated with a closed form from the following estimating equation,

$$U(\beta) = \sum_{i=1}^n \int_0^\infty \left(B_i, W_i^T(u)\right)^T \{dN_i(t) - Y_i(t)d\hat{\Lambda}_0(\beta,t) - Y_i(t)\beta^T \left(B_i, W_i^T(u)\right)^T dt\},$$

which is equivalent to

$$U(\beta) = \sum_{i=1}^n \int_0^\infty \{\left(B_i, W_i^T(u)\right)^T - \left(\bar{B}, \bar{W}^T(u)\right)^T\}\{dN_i(t) - Y_i(t)\beta^T \left(B_i, W_i^T(u)\right)^T dt\}, \tag{5.4}$$

where

$$\bar{B} = \sum_{i=1}^n \frac{Y_i(t)}{\sum_{i=1}^n Y_i(t)} B_i,$$
$$\bar{W}(t) = \sum_{i=1}^n \frac{Y_i(t)}{\sum_{i=1}^n Y_i(t)} W_i(t).$$

Introduce the following notations,

$$\bar{Y}(t) = \frac{1}{n}\sum_{i=1}^{n} Y_i(t), \quad C(t) = \int_0^t e_0 du = e_0 t$$
$$G(t) = \int_0^t \{z_0(s) - \bar{Z}(s)\} ds,$$
$$\Sigma = \left[\frac{1}{n}\sum_{i=1}^{n}\int_0^{\infty} \{Z_i(t) - \bar{Z}(t)\}^{\otimes 2} Y_i(t) dt\right]',$$

where $z_0(t) = \left(e_0, w_0^T(t)\right)^T$ is a chosen level of covariates.

We propose the confidence interval for $Q(t) = \sqrt{n}\left(\hat{A}(t) - A(t)\right)$ using theorems for empirical processes. The limiting distribution is proved in the future.

Chapter 6

FUTURE WORK

As mentioned above, we have studied the comparison of survival functions using counting process procedures as well as asymptotic properties of martingales. Intensive simulation results conducted insofar have produced statistically satisfactory outcomes when the sample size is relatively large. However, in small sample cases, especially when less than 50, asymptotic approximations start to exhibit reduced performance. To overcome these limitations of the normal approximation and improve the coverage accuracy of the corresponding confidence bands, we are trying to employ the empirical likelihood method to derive simultaneous confidence bands for the ratio and if possible, difference, allowing covariate-adjusted survival functions for the treatment comparison between two populations.

Apart from the normal approximation, empirical likelihood based confidence band has excellent coverage accuracy in small samples, as well as various desirable properties including, but not limited to, range-preserving, transform-respecting and asymmetric since it relies solely on the features of the data to determine its shape. It is also worth mentioning that empirical likelihood based bands are easier to construct without much complication of deriving a variance estimator. This is particular important in two-sample applications when quite frequently, covariance formulation becomes problematic.

Moreover, from the normal approximation point of view, due to the modeling limitations of the proportional hazard model, the transformation model has gained a lot of attentions recently in the survival literatures. Chen et al. [15] developed a transformation model for the attributable fraction function, which almost completed any work that might be relevant. We are trying to develop an alternative approach using the additive risk model. Brief methodological structure is given in Chapter 5.

Appendix A

PROOFS OF THEOREMS

Write $\varepsilon_n = n^{-s}$, where $s \in (1/3, 1/2)$, and $n = n_1 + n_2$. Note that $\eta_0 = S_2(t)$.

Lemma 1. *Under the conditions of Theorem 1, if η satisfies $|\eta - \eta_0| \leq \varepsilon_n$, then the Lagrange multipliers (λ_1, λ_2) of equation (2.3) and (2.4) satisfy*

$$\frac{\lambda_j}{n} = O(\varepsilon_n), \quad j = 1, 2, \quad a.s. \tag{A.1}$$

uniformly for $t \in [\tau_1, \tau_2]$.

Proof. Adopting the notations used in Shen and He [71], write

$$\begin{aligned} A_1(\eta, t) &= \ln\left(\frac{\eta}{\eta + \theta_0 - \theta_0\eta}\right) - \ln \hat{S}_1(t), \\ A_2(\eta, t) &= \ln(\eta) - \ln \hat{S}_2(t). \end{aligned} \tag{A.2}$$

Applying the LIL of $\hat{S}_1(t)$ Csörgö and Horváth [19], we know by Taylor expansion that there exists a η^* between η_0 and η such that

$$\begin{aligned} A_1(\eta, t) &= \ln(\theta_0(t) + \eta) - \ln(\theta_0(t) + \eta_0) + \ln(\theta_0(t), \eta_0) - \ln(\hat{S}_1(t)) \\ &= \frac{\theta_0(t)(\eta - \eta_0)}{\eta^*(\eta^* + \theta_0(t) - \theta_0(t)\eta^*)} + \ln S_1(t) - \ln \hat{S}_1(t) \\ &= O(\varepsilon_n) + O((n_1/\ln n_1)^{-1/2}) = O(\varepsilon_n), \quad \text{a.s.} \end{aligned} \tag{A.3}$$

uniformly for $t \in [\tau_1, \tau_2]$.

Then, by equation (4.3) of Shen and He [71], we know

$$\begin{aligned} \lambda_1 A_1(\eta, t) &= |\lambda_1| \sum_{i=1}^{K_1(t)} \left| \ln\left(1 - \frac{d_{1i}}{r_{1i} + \lambda_1}\right) - \ln\left(1 - \frac{d_{1i}}{r_{1i}}\right) \right| \\ &\geq \frac{\lambda_1^2}{n_1 + |\lambda_1| \max_{i:T_{1i} \leq t}\{|n_1/r_{1i}|\}} \cdot \tilde{\sigma}_1^2(t), \end{aligned} \tag{A.4}$$

where

$$\tilde{\sigma}_1^2(t) = n_1 \sum_{i=1}^{K_1(t)} \frac{d_{1i}}{r_i^2} = \int_0^t \frac{\mathrm{d}F_1(u)}{\bar{F}_1(u-)\bar{H}_1(u-)} + o(1) \geq \sigma_1^2(\tau_1)/2. \quad \text{a.s.} \tag{A.5}$$

By the strong law of large numbers plus the monotonicity of $\bar{H}_1(t)$, we know

$$\max_{i:T_{1i}\leq t} \left| \frac{n_1}{r_{1i}} \right| = \frac{1}{\bar{H}_1(t)} \leq \frac{2}{\bar{H}_1(\tau_2)} \tag{A.6}$$

Thus, combining (A.4)-(A.6), for $t \in [\tau_1, \tau_2]$, we get almost surely that for large n,

$$|A_1(\eta, t)| \geq \frac{|\lambda_1|}{n_1 + 2|\lambda_1|\bar{H}_1^{-1}(\tau_2)} \cdot \frac{\sigma_1^2(\tau_1)}{2}. \tag{A.7}$$

Plugging (A.7) into (A.3), it is easy to see that

$$\lambda_1/n_1 = O(\varepsilon_n), \quad \text{a.s.}$$

uniformly for $t \in [\tau_1, \tau_2]$.

Similarly, it can also be proved that

$$\lambda_2/n_2 = O(\varepsilon_n) \text{ a.s. uniformly for } t \in [\tau_1, \tau_2].$$

Hence proves the lemma. □

Lemma 2. *Under the conditions of Theorem 1, for large n, there exists almost surely a solution to equation (2.5), denoted as $\eta_E(t)$, such that $\mathcal{R}(\theta_0, \eta, t)$ attains its maximum value at $\eta = \eta_E(t)$, and*

$$\eta_E(t) \to \eta_0 = S_2(t), \quad a.s. \tag{A.8}$$

as $n \to \infty$.

Proof. First, for (j, i) such that $X_{ji} < \tau_2$, we know almost surely that,

$$\frac{n_j}{r_{ji}} \leq \frac{n_j}{\sum_{k=1}^{n_j}(X_{jk} \geq \tau_2)} \leq \frac{2}{\bar{H}_j(\tau_2)}.$$

From Shen and He [71], we have

$$\ln\left(1-\frac{d_{ji}}{r_{ji}+\lambda_j}\right)=\ln\left(1-\frac{d_{ji}}{r_{ji}}\right)+\frac{d_{ji}}{r_{ji}(r_{ji}-d_{ji})}\lambda_j+O\left(\frac{\varepsilon_n^2}{n}\right). \tag{A.9}$$

Write $\eta_n=\eta_0+\varepsilon_n$. Using equation (2.3) and (2.4), equation (A.9) yields that

$$\begin{aligned}\ln\left(\frac{\eta}{\eta+\theta_0-\theta_0\eta}\right)&=\ln\hat{S}_1(t)+\frac{\lambda_1\hat{\sigma}_1^2(t)}{n_1}+O(\varepsilon_n^2)\ \text{ a.s.}\\ \ln(\eta)&=\ln\hat{S}_2(t)+\frac{\lambda_2\hat{\sigma}_2^2(t)}{n_2}+O(\varepsilon_n^2)\ \text{ a.s.}\end{aligned}$$

Thus,

$$\lambda_j(\eta,t)=\frac{n_j}{\hat{\sigma}_j^2(t)}A_j(\eta,t)+O(n_j\varepsilon_n^2),\ \text{ a.s.},\quad j=1,2. \tag{A.10}$$

By Csörgö and Horváth [19], it is easy to see that almost surely,

$$A_j(\eta_0,t)=\ln S_j(t)-\ln\hat{S}_j(t)=o(\varepsilon_n)\quad j=1,2. \tag{A.11}$$

Similar to equation (4.12) of Shen and He [71], by Taylor expansion, write

$$\begin{aligned}&-2\ln\mathcal{R}(\theta_0,\eta_n,t)\\&\quad=\frac{n_1}{\hat{\sigma}_1^2(t)}\left(A_1(\eta_0,t)+\frac{\theta_0(t)\varepsilon_n}{\eta_1(\theta_0(t)+\eta_1-\theta_0(t)\eta_1)}\right)^2\\&\quad+\frac{n_2}{\hat{\sigma}_2^2(t)}\left(A_2(\eta_0,t)+\frac{\varepsilon_n}{\eta_2}\right)^2+O(n\varepsilon_n^3),\quad\text{a.s.}\end{aligned} \tag{A.12}$$

where η_1 and η_2 are all between η_0 and η_n.

Therefore, plugging (A.10) and (A.11) into (A.12), for sufficiently large enough n, we can show that almost surely

$$-2\ln\mathcal{R}(\theta_0,\eta_n,t)\geq\varepsilon_n^2\sum_{j=1}^{2}\frac{n_j}{2\sigma_1^2(t)S_j^2(t)}.$$

On the other hand, it can be shown in the same sense that

$$\begin{aligned} -2\ln\mathcal{R}(\theta_0,\eta_0,t) &= \sum_{j=1}^{2}\frac{n_j}{\hat{\sigma}_j^2(t)}A_j(\eta_0,t)^2 + O(n\varepsilon_n^3) \\ &= o(n\varepsilon_n^2) \quad \text{a.s.} \end{aligned}$$

Thus, for n large enough,

$$-2\ln\mathcal{R}(\theta_0,\eta_0+\varepsilon_n,t) > -2\ln\mathcal{R}(\theta_0,\eta_0,t) \quad \text{a.s.}$$

Similarly,

$$-2\ln\mathcal{R}(\theta_0,\eta_0-\varepsilon_n,t) > -2\ln\mathcal{R}(\theta_0,\eta_0,t) \quad \text{a.s.}$$

This means $-2\ln\mathcal{R}(\theta_0,\eta,t)$ attains its minimum in $(\eta_0-\varepsilon_n,\eta_0+\varepsilon_n)$. Hence, there exists $\eta_E = \arg\max_{\eta}\mathcal{R}(\theta_0,\eta,t)$ in $(\eta_0-\varepsilon_n,\eta_0+\varepsilon_n)$ that satisfies equation (A.8). □

Proof of Theorem 1. Let $\lambda_1 = n_i\gamma_i$, $i = 1,2$ and plug-in into the left-hand sides of equation (2.3)-(2.5). Then, denote the three equations as $E_j(\eta,\gamma_1,\gamma_2,t)$, $j = 1,2,3$, respectively. We calculate

$$\begin{aligned} \hat{J}(\eta,t) &= \left.\frac{\partial E_1,E_2,E_3}{\partial(\eta,\gamma_1,\gamma_2)}\right|_{(\eta,\gamma_1,\gamma_2,t)=(\eta,0,0,t)} \\ &= \begin{pmatrix} -\frac{\theta_0(t)}{\eta(\eta+\theta_0(t)-\theta_0(t)\eta)} & \hat{\sigma}_1^2(t) & 0 \\ -1/\eta & 0 & \hat{\sigma}_2^2(t) \\ 0 & \hat{p}_1\frac{\theta_0(t)}{\eta+\theta_0(t)-\theta_0(t)\eta} & \hat{p}_2 \end{pmatrix}, \end{aligned}$$

where $\hat{p}_j = n_j/n$, $j = 1, 2$.

Denote $\gamma_{jE} = \gamma_j(\eta_E, t)$, $j = 1, 2$. By the Taylor expansion, we have

$$\begin{aligned}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} E_1(\eta_E, \gamma_{1E}, \gamma_{2E}, t) \\ E_2(\eta_E, \gamma_{1E}, \gamma_{2E}, t) \\ E_3(\eta_E, \gamma_{1E}, \gamma_{2E}, t) \end{pmatrix} \\ &= \begin{pmatrix} E_1(\eta_0, 0, 0, t) \\ E_2(\eta_0, 0, 0, t) \\ E_3(\eta_0, 0, 0, t) \end{pmatrix} + \hat{J}(\eta_0, t) \begin{pmatrix} \eta_E - \eta_0 \\ \gamma_{1E} \\ \gamma_{2E} \end{pmatrix} + O_p(\varepsilon_n^2).\end{aligned}$$

Note that $E_i(\eta_0, 0, 0, t) = \ln \hat{S}_i(t) - \ln S_i(t)$, $i = 1, 2$, $E_3(\eta_0, 0, 0, t) = 0$ and $\varepsilon_n^2 = o(n^{-1/2})$.

Therefore,

$$\begin{aligned}\begin{pmatrix} \eta_E - \eta_0 \\ \gamma_{1E} \\ \gamma_{2E} \end{pmatrix} &= -\hat{J}^{-1}(\eta_0, t) \begin{pmatrix} \ln \hat{S}_1(t) - \ln S_1(t) \\ \ln \hat{S}_2(t) - \ln S_2(t) \\ 0 \end{pmatrix} + o_p(n^{-1/2}) \\ &= \frac{\ln \hat{S}_1(t) - \ln S_1(t)}{\det(\hat{J}(\eta_0, t))} \begin{pmatrix} \hat{p}_1 \hat{\sigma}_2^2(t) \theta_0(t) (\eta_0 + \theta_0(t) - \theta_0(t)\eta_0)^{-1} \\ -\hat{p}_2 \eta_0^{-1} \\ \hat{p}_1 \theta_0(t) \eta_0^{-1} (\eta_0 + \theta_0(t) - \eta_0 \theta_0(t))^{-1} \end{pmatrix} \\ &\quad + \frac{\ln \hat{S}_2(t) - \ln S_2(t)}{\det(\hat{J}(\eta_0, t))} \begin{pmatrix} \hat{p}_2 \hat{\sigma}_1^2(t) \\ \hat{p}_2 \theta_0(t) \eta_0^{-1} (\eta_0 + \theta_0(t) - \theta_0(t)\eta_0)^{-1} \\ -\hat{p}_1 \theta_0^2(t) \eta_0^{-1} (\eta_0 + \theta_0(t) - \theta_0(t)\eta_0)^{-2} \end{pmatrix} + o_p(n^{-1/2}), \quad \text{(A.13)}\end{aligned}$$

where

$$\det(\hat{J}(\eta, t)) = \frac{\hat{\sigma}_1^2(t)}{\eta} \hat{p}_2 + \frac{\hat{\sigma}_2^2(t) \theta_0^2(t)}{\eta(\eta + \theta_0(t) - \theta_0(t)\eta)^2} \hat{p}_1.$$

By equation (2.11), we can show that

$$\hat{J}(\eta, t) \xrightarrow{P} J(\eta, t) = \begin{pmatrix} -\frac{\theta_0(t)}{\eta(\eta + \theta_0(t) - \theta_0(t)\eta)} & \sigma_1^2(t) & 0 \\ -1/\eta & 0 & \sigma_2^2(t) \\ 0 & p_1 \frac{\theta_0(t)}{\eta + \theta_0(t) - \theta_0(t)\eta} & p_2 \end{pmatrix}, \quad \text{(A.14)}$$

as $n \to \infty$.

By Lemma 1 and Lemma 4.1 of Shen and He [71], we know

$$\begin{aligned}
\frac{\lambda_1^2(\eta_E,\cdot)}{n_1} &= (\sqrt{n_1}\gamma_1(\eta_E,\cdot))^2 \\
&= \left(-\frac{\sqrt{n_1}(\ln\hat{S}_1(\cdot)-\ln S_1(\cdot))}{\det(\hat{J}(\eta_0,\cdot))}\cdot\frac{\hat{p}_2}{\eta_0^2}+\frac{\sqrt{n_2}(\ln\hat{S}_2(\cdot)-\ln S_2(\cdot))}{\det(\hat{J}(\eta_0,\cdot))}\cdot\frac{\sqrt{n_1}\hat{p}_2\theta_0}{\sqrt{n_2}\eta_0^2(\eta_0+\theta_0-\eta_0\theta_0)}\right)^2 \\
&\xrightarrow{\mathfrak{D}} \frac{p_1p_2^2}{\det(J(\eta_0,\cdot))^2\eta_0^2(1-\eta_0)^2}\left(-\frac{W_1(\sigma_1^2(\cdot))}{\sqrt{p_1}}+\frac{\theta_0}{\eta_0+\theta_0-\theta_0\eta_0}\cdot\frac{W_2(\sigma_2^2(\cdot))}{\sqrt{p_2}}\right)^2 \\
&\overset{\mathfrak{D}}{=} \frac{p_1p_2^2}{\det(J(\eta_0,\cdot))^2\eta_0^2(1-\eta_0)^2}\left(\frac{(1-S_2(\cdot))W_1(\sigma_1^2(\cdot))}{\sqrt{p_1}}+\frac{(1-S_1(\cdot))W_2(\sigma_2^2(\cdot))}{\sqrt{p_2}}\right)^2. \qquad \text{(A.15)}
\end{aligned}$$

From (2.5), we know

$$\frac{\lambda_1\theta_0(t)}{\eta_E(\eta_E+\theta_0(t)-\eta_E\theta_0(t))}+\frac{\lambda_2}{\eta_E}=0.$$

It yields from (A.12) that

$$\begin{aligned}
-2\ln\mathcal{R}(\theta_0(t),\eta_E,t) &= \sum_{j=1}^{2}\frac{\lambda_j^2(\eta_E,t)}{n_j}\hat{\sigma}_j^2(t)+o_p(1) \\
&= \frac{\lambda_1(\eta_E,t)}{n_1}\left(\hat{\sigma}_1^2(t)+\frac{\hat{p}_1}{\hat{p}_2}\frac{\theta_0^2(t)}{(\eta_E+\theta_0(t)-\eta_E\theta_0(t))^2}\hat{\sigma}_2^2(t)\right)+o_p(1) \\
&= \frac{\lambda_1^2(\eta_E,t)}{n_1}\frac{\det(\hat{J}(\eta_E,t))\eta_E^2}{\hat{p}_2}+o_p(1).
\end{aligned}$$

Hence, from Lemma 1, equation (A.14) and (A.15), it is easy to see that

$$\begin{aligned}
-2\ln\mathcal{R}(\theta_0,\eta_E,\cdot) &\xrightarrow{\mathfrak{D}} \frac{p_1p_2}{\det(J(\eta_0,\cdot))\eta_0(1-\eta_0)^2}W^2(\sigma^2(\cdot)) \\
&= \frac{1}{\sigma^2(\cdot)}\left(\frac{(1-S_2(\cdot))W_1(\sigma_1^2(\cdot))}{\sqrt{p_1}}+\frac{(1-S_1(\cdot))W_2(\sigma_2^2(\cdot))}{\sqrt{p_2}}\right)^2. \quad \square
\end{aligned}$$

Proof of Theorem 2. Note that by the Taylor expansion,

$$\begin{aligned}
n^{1/2}\left\{\hat{\Lambda}_i(t;z_0)-\Lambda_i(t;z_0)\right\} &= n^{1/2}\left\{e^{\hat{\beta}'z_0(t)}\hat{\Lambda}_{i0}(t;z_0)-e^{\beta'z_0(t)}\Lambda_{i0}(t;z_0)\right\} \\
&\approx n^{1/2}\left[e^{\beta'z_0(t)}\hat{\Lambda}_{i0}(t;z_0)+e^{\beta'z_0(t)}z_0(t)(\hat{\beta}-\beta_0)\hat{\Lambda}_{i0}(t;z_0)-e^{\beta'z_0(t)}\Lambda_{i0}(t;z_0)\right] \\
&= n^{1/2}\left\{e^{\beta'z_0(t)}[\hat{\Lambda}_{i0}(t;z_0)-\Lambda_{i0}(t;z_0)]\right\}+\{e^{\beta'z_0(t)}z_0(t)\hat{\Lambda}_{i0}(t;z_0)\}[n^{1/2}(\hat{\beta}-\beta)]. \qquad \text{(A.16)}
\end{aligned}$$

Corollary 3.5 of Andersen and Gill [4] has shown that by the Taylor expansion

$$\begin{aligned} & n^{1/2}\left(\hat{\Lambda}_{i0}(t)-\Lambda_{i0}(t)\right) \\ & \quad \overset{\mathfrak{D}}{\approx} \int_0^t \frac{n^{1/2}d\bar{M}_i(u)}{\sum_{j=1}^{n_i} Y_{ij}(u)e^{\beta' Z_{ij}(u)}} + \left[-\int_0^t \frac{S_i^{(1)}(\beta,u)}{S_i^{(0)}(\beta,u)^2} d\bar{N}_i(u)\right]' \left\{n^{1/2}(\hat{\beta}-\beta)\right\} \\ & \quad \overset{\mathfrak{D}}{=} \int_0^t \frac{n^{1/2}d\bar{M}_i(u)}{\sum_{j=1}^{n_i} Y_{ij}(u)e^{\beta' Z_{ij}(u)}} + \left[-\int_0^t e(\beta,u)d\Lambda_{i0}(u)\right]' \left\{n^{1/2}(\hat{\beta}-\beta)\right\}. \end{aligned} \tag{A.17}$$

Moreover, from Theorem 3.2 of Andersen and Gill [4], it is easy to see that

$$\begin{aligned} n^{1/2}(\hat{\beta}-\beta) &= \{n^{-1}(\mathcal{I}(\beta^*,\infty))\}^{-1}[n^{-1/2}U(\beta,\infty)] \\ &= \sum_{i=1}^{2}\frac{n_i}{n}\int_0^\infty V_i(\beta^*,t)\frac{d\bar{N}_i(t)}{n_i}\cdot\left[n^{-1/2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(u)-E_i(\beta_0,u)\}dM_{ij}(u)\right] \\ &\overset{\mathbf{P}}{\to} \Sigma^{-1}\left[n^{-1/2}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty n^{-1/2}\{Z_{ij}(u)-E_i(\beta_0,u)\}dM_{ij}(u)\right], \end{aligned} \tag{A.18}$$

where $\Sigma=\sum_{i=1}^{2}\int_0^\infty v_i(\beta,t)s_i^{(0)}(\beta,t)d\Lambda_{i0}(t)$ is the covariance matrix.
Therefore,

$$W(t;z_0)=n^{1/2}[(\hat{\Lambda}_2(t;z_0)-\hat{\Lambda}_1(t;z_0))-(\Lambda_2(t;z_0)-\Lambda_1(t;z_0))],$$

is asymptotically equivalent to

$$\begin{aligned} \widetilde{W}(t;z_0) =& \frac{1}{\sqrt{n}}\frac{1}{p_2}\int_0^t \frac{e^{\beta_0' z_0(u)}d\bar{M}_2(u)}{S_2^{(0)}(\beta_0,u)} - \frac{1}{\sqrt{n}}\frac{1}{p_1}\int_0^t \frac{e^{\beta_0' z_0(u)}d\bar{M}_1(u)}{S_1^{(0)}(\beta_0,u)} \\ & + (h_2(t;z_0)-h_1(t;z_0))'\Sigma^{-1}\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(u)-E_i(\beta_0,u)\}dM_{ij}(u)\right]. \end{aligned}$$

Since $\widetilde{W}(t;z_0)$ is a martingale, by Rebolledo's martingale central limit theorem, we can show that $\widetilde{W}(t;z_0)$ converges weakly to a zero mean Gaussian martingale on $[0,\tau]$, where $\tau<\inf_{t>0}\{t:$

$E(Y_{ij}(t)) = 0\}$, with covariance function being

$$\xi(t, v; z_0) = \sum_{i=1}^{2} \frac{1}{p_i} \int_0^{t\wedge v} \frac{e^{2\beta_0' z_0(u)} d\bar{N}_i(u)}{n_i \left[s_i^{(0)}(\beta_0, u) \right]^2} + (h_1(t; z_0) - h_2(t; z_0))' \Sigma^{-1} (h_1(v; z_0) - h_2(v; z_0)). \quad \square \tag{A.19}$$

Proof of Theorem 3. The limiting distribution of $W_{OR}(t; z_0)$ can be derived in a fairly analogous manner. Thus, we only give a brief explanation.

First, consider taking a logarithm. Using the functional delta method, it is easy to see that

$$\begin{aligned}
& n^{1/2} \left[\log \left(\frac{\hat{S}_1(t; z_0)}{1 - \hat{S}_1(t; z_0)} \Big/ \frac{\hat{S}_2(t; z_0)}{1 - \hat{S}_2(t; z_0)} \right) - \log \left(\frac{S_1(t; z_0)}{1 - S_1(t; z_0)} \Big/ \frac{S_2(t; z_0)}{1 - S_2(t; z_0)} \right) \right] \\
&= n^{1/2} \left[\log \left(\frac{e^{\hat{\Lambda}_2(t; z_0)} - 1}{e^{\hat{\Lambda}_1(t; z_0)} - 1} \right) - \log \left(\frac{e^{\Lambda_2(t; z_0)} - 1}{e^{\Lambda_1(t; z_0)} - 1} \right) \right] \\
&\stackrel{\mathfrak{D}}{=} n^{1/2} \left\{ \frac{1}{1 - S_2(t; z_0)} [\hat{\Lambda}_2(t; z_0) - \Lambda_2(t; z_0)] - \frac{1}{1 - S_1(t; z_0)} [\hat{\Lambda}_1(t; z_0) - \Lambda_1(t; z_0)] \right\}.
\end{aligned}$$

It yields that

$$\begin{aligned}
W_{OR}(t) =& n^{1/2} \left[\frac{\hat{S}_1(t; z_0)}{1 - \hat{S}_1(t; z_0)} \Big/ \frac{\hat{S}_2(t; z_0)}{1 - \hat{S}_2(t; z_0)} - \frac{S_1(t; z_0)}{1 - S_1(t; z_0)} \Big/ \frac{S_2(t; z_0)}{1 - S_2(t; z_0)} \right] \\
\stackrel{\mathfrak{D}}{=} & \sqrt{n} \left[\frac{S_1(t; z_0)}{1 - S_1(t; z_0)} \Big/ \frac{S_2(t; z_0)}{1 - S_2(t; z_0)} \right] \cdot \\
& \left\{ \frac{1}{1 - S_2(t; z_0)} [\hat{\Lambda}_2(t; z_0) - \Lambda_2(t; z_0)] - \frac{1}{1 - S_1(t; z_0)} [\hat{\Lambda}_1(t; z_0) - \Lambda_1(t; z_0)] \right\}.
\end{aligned} \tag{A.20}$$

By Andersen and Gill [4], (A.20) is asymptotically equivalent to

$$\widetilde{W}_{OR}(t;z_0) = \left[\frac{S_1(t;z_0)}{1-S_1(t;z_0)}\Big/\frac{S_2(t;z_0)}{1-S_2(t;z_0)}\right]\cdot \left\{\frac{1}{\sqrt{n}}\frac{1}{p_2}\frac{1}{1-S_2(t;z_0)}\int_0^t \frac{e^{\beta_0' z_0(u)}d\bar{M}_2(u)}{S_2^{(0)}(\beta_0,u)} - \frac{1}{\sqrt{n}}\frac{1}{p_1}\frac{1}{1-S_1(t;z_0)}\int_0^t \frac{e^{\beta_0' z_0(u)}d\bar{M}_1(u)}{S_1^{(0)}(\beta_0,u)} + \left(\frac{h_2(t;z_0)}{1-S_2(t;z_0)} - \frac{h_1(t;z_0)}{1-S_1(t;z_0)}\right)'\Sigma^{-1} \left[\frac{1}{\sqrt{n}}\sum_{i=i}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(u)-E_i(\beta_0,u)\}dM_{ij}(u)\right]\right\}.$$

Similar to (A.19), the covariate function of $\widetilde{W}_{OR}(t;z_0)$ is given by

$$\xi_{OR}(t;z_0) = \left[\frac{S_1(t;z_0)}{1-S_1(t;z_0)}\Big/\frac{S_2(t;z_0)}{1-S_2(t;z_0)}\cdot\frac{S_1(v;z_0)}{1-S_1(v;z_0)}\Big/\frac{S_2(v;z_0)}{1-S_2(v;z_0)}\right]\cdot \left\{\sum_{i=1}^{2}\frac{1}{p_i}\frac{1}{1-S_i(t;z_0)}\frac{1}{1-S_i(v;z_0)}\int_0^{t\wedge v}\frac{e^{2\beta_0' z_0(u)}d\bar{N}_i(u)}{n_i\left[s_i^{(0)}(\beta_0,u)\right]^2} + \left(\frac{h_1(t;z_0)}{1-S_1(t;z_0)} - \frac{h_2(t;z_0)}{1-S_2(t;z_0)}\right)'\Sigma^{-1} \left(\frac{h_1(v;z_0)}{1-S_1(v;z_0)} - \frac{h_2(v;z_0)}{1-S_2(v;z_0)}\right)\right\}. \quad \square \tag{A.21}$$

Proof of Theorem 4. Note that by (4.4), we have

$$\hat{\Lambda}_i(t;z_0) - \Lambda_i(t;z_0) = \sum_{j=1}^{n_i}\int_0^t \frac{1}{n_i\bar{Y}_i(u)}dM_{ij}(u) + (\hat{\beta}-\beta)'\int_0^t\{z_0(u)-\bar{Z}_i(u)\}du. \tag{A.22}$$

Moreover, by the Taylor expansion of $U(\hat{\beta})$ at β, it is easy to see that

$$(\hat{\beta}-\beta)' = \frac{\Sigma^{-1}}{n}\left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty\{Z_{ij}(t)-\bar{Z}_i(t)\}dM_{ij}(t)\right]'. \tag{A.23}$$

Taking (A.23) back into (A.22) yields that $L(t;z_0)$ is asymptotically equivalent to

$$\begin{aligned}\widetilde{L}(t;z_0) =& \frac{1}{\sqrt{n}}\frac{1}{p_2}\sum_{j=1}^{n_2}\int_0^t \frac{1}{\bar{Y}_2(s)}dM_{2j}(s) - \frac{1}{\sqrt{n}}\frac{1}{p_1}\sum_{j=1}^{n_1}\int_0^t \frac{1}{\bar{Y}_1(s)}dM_{1j}(s) \\ &+ \frac{G_2'(t;z_0)-G_1'(t;z_0)}{\sqrt{n}}\Sigma^{-1}\left[\sum_{i=1}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(s)-\bar{Z}_i(s)\}dM_{ij}(s)\right],\end{aligned}$$

where $p_i = \lim n_i/n$.

Write

$$G(t;z_0) = G_2(t;z_0) - G_1(t;z_0),$$

and

$$\tau = \inf\{t \geq 0; H_1(t) = H_2(t) = 1\},$$

where $H_i(t)$ is the distriubtion function of the observed failure time X_{ij}. Notice that by Theorem 2 of Song et al. [74], $G(t;z_0)$ and Σ^{-1} all converges to some nonrandom functions and

$$\frac{1}{\sqrt{n}}\sum_{i=i}^{2}\sum_{j=1}^{n_i}\int_0^\infty \{Z_{ij}(s)-\bar{Z}_i(s)\}dM_{ij}(s)$$

converges in distribution, if $\lim n_i/n = p_i \in (0,1)$. Therefore, we know $\widetilde{L}(t;z_0)$ is tight. Moreover, similar to equation (2.2) of Song et al. [74], $\widetilde{L}(t;z_0)$ is indeed a martingale, and thus by the Linderberg-Feller theorem and the above tightness, we know that the process $\widetilde{L}(t;z_0)$ converges weakly to a zero mean Gaussian process on $[0,\tau)$. The weak convergence of $\widetilde{L}_{OR}(t;z_0)$ can be proved analogously.

For covariance formulae, note that $\mathrm{E}[M_{ij}(u)] = 0$ and $\mathrm{var}[M_{ij}(u)] = \mathrm{E}[N_{ij}(u)]$, thus it follows from Chapter II3.2 of Andersen et al. [5] that the covariance matrix of $\widetilde{L}(t;z_0)$ can be consistently estimated by

$$\begin{aligned}\xi_L(t;z_0) =& \sum_{i=1}^{2}\left[\frac{1}{n}\frac{1}{p_i^2}\sum_{j=1}^{n_i}\int_0^{t\wedge s}\frac{dN_{ij}(u)}{(\bar{Y}_i(u))^2} + (-1)^i\frac{G_R'(s;z_0)}{p_i}\Sigma^{-1}D_{i1}(t) + (-1)^i\frac{G_R'(t;z_0)}{p_i}\Sigma^{-1}D_{i1}(s)\right] \\ &+ G_R'(t;z_0)\Sigma^{-1}D_{i2}(\Sigma^{-1})'G_R(s;z_0),\end{aligned}$$

where

$$D_{i1}(t) = \sum_{j=1}^{n_i} \int_0^t \frac{\{Z_{ij}(u) - \bar{Z}_i(u)\} dN_{ij}(u)}{\sum_{k=1}^{n_i} Y_{ik}(u)}, \tag{A.24}$$

$$D_{i2} = \frac{1}{n} \sum_{j=1}^{n_i} \int_0^\infty \{Z_{ij}(t) - \bar{Z}_i(t)\}^{\otimes 2} dN_{ij}(t), \tag{A.25}$$

and for column vector a, $a^{\otimes 2}$ denotes the outer product of a. $\square$

Proof of Theorem 5. Now, we consider the odds ratio of survival functions. Denote $\theta(t; z_0) = \frac{1-S_1(t;z_0)}{S_1(t;z_0)} \Big/ \frac{1-S_2(t;z_0)}{S_2(t;z_0)}$. We know by simple algebra

$$\theta(t; z_0) = \frac{e^{\Lambda_1(t;z_0)} - 1}{e^{\Lambda_2(t;z_0)} - 1}.$$

From (A.20), we know

$$\begin{aligned} L_{OR}(t) &= n^{1/2} \left[\frac{1 - \hat{S}_1(t; z_0)}{\hat{S}_1(t; z_0)} \Big/ \frac{1 - \hat{S}_2(t; z_0)}{\hat{S}_2(t; z_0)} - \frac{1 - S_1(t; z_0)}{S_1(t; z_0)} \Big/ \frac{1 - S_2(t; z_0)}{S_2(t; z_0)} \right] \\ &\stackrel{\mathfrak{D}}{=} \sqrt{n}\, \frac{(1 - S_2(t; z_0)) S_1(t; z_0)}{S_2(t; z_0)(1 - S_1(t; z_0))} \cdot \left\{ \frac{1}{1 - S_2(t; z_0)} [\hat{\Lambda}_2(t; z_0) - \Lambda_2(t; z_0)] \right. \\ &\qquad \left. - \frac{1}{1 - S_1(t; z_0)} [\hat{\Lambda}_1(t; z_0) - \Lambda_1(t; z_0)] \right\}. \end{aligned} \tag{A.26}$$

By Andersen and Gill [4], (A.26) is asymptotically equivalent to

$$\begin{aligned} \widetilde{L}_{OR}(t; z_0) =& \frac{(1 - S_2(t; z_0)) S_1(t; z_0)}{S_2(t; z_0)(1 - S_1(t; z_0))} \cdot \\ & \left\{ \frac{1}{\sqrt{n}} \frac{1}{p_2(1 - S_2(t; z_0))} \sum_{j=1}^{n_2} \int_0^t \frac{1}{\bar{Y}_2(s)} dM_{2j}(s) \right. \\ & - \frac{1}{\sqrt{n}} \frac{1}{p_1(1 - S_1(t; z_0))} \sum_{j=1}^{n_1} \int_0^t \frac{1}{\bar{Y}_1(s)} dM_{1j}(s) \\ & + \frac{1}{\sqrt{n}} \left[\frac{G_2'(t; z_0)}{1 - S_2(t; z_0)} - \frac{G_1'(t; z_0)}{1 - S_1(t; z_0)} \right] \Sigma^{-1} \\ & \left. \left[\sum_{i=1}^{2} \sum_{j=1}^{n_i} \int_0^\infty \{Z_{ij}(s) - \bar{Z}_i(s)\} dM_{ij}(s) \right] \right\}. \end{aligned} \tag{A.27}$$

Under a similar logic, write

$$G_{OR}(t;z_0) = \frac{G_2(t;z_0)}{1-S_2(t;z_0)} - \frac{G_1(t;z_0)}{1-S_1(t;z_0)},$$

we are able to derive a consistent estimator for the covariance function of $\widetilde{L}_{OR}(t;z_0)$ as

$$\begin{aligned}\xi_{L_{OR}}(t;z_0) = \frac{(1-S_2(t;z_0))S_1(t;z_0)}{S_2(t;z_0)(1-S_1(t;z_0))}\frac{(1-S_2(s;z_0))S_1(s;z_0)}{S_2(s;z_0)(1-S_1(s;z_0))}\cdot \\ \left\{\sum_{i=1}^{2}\left[\frac{1}{n}\frac{1}{p_i^2}\frac{1}{(1-S_i(t;z_0))}\frac{1}{(1-S_i(s;z_0))}\sum_{j=1}^{n_i}\int_0^{t\wedge s}\frac{dN_{ij}(u)}{(\bar{Y}_i(u))^2}\right.\right. \\ + (-1)^i\frac{G_R'(s;z_0)}{(1-S_i(t;z_0))p_i}\Sigma^{-1}D_{i1}(t) \\ \left.\left. + (-1)^i\frac{G_R'(t;z_0)}{p_i(1-S_i(s;z_0))}\Sigma^{-1}D_{i1}(s)\right] \right. \\ \left. + G_R'(t;z_0)\Sigma^{-1}D_{i2}(\Sigma^{-1})'G_R(s;z_0)\right\},\end{aligned}$$

where $D_{i1}(t)$ and D_{i2} are specified in (A.24) and (A.25). $\square$

-

Appendix B

TABLES

Table 1: Coverage probability of simultaneous confidence bands for $\theta_0(t),\ t \in [0.1, 2.5]$

CR	n_1	$\alpha = 0.05$		$\alpha = 0.10$	
		EL	NA	EL	NA
0.10	30	0.936	0.886	0.887	0.873
	50	0.949	0.912	0.89	0.889
	80	0.950	0.945	0.892	0.904
	100	0.965	0.949	0.926	0.890
0.30	30	0.871	0.860	0.899	0.833
	50	0.946	0.924	0.900	0.895
	80	0.949	0.943	0.901	0.919
	100	0.960	0.950	0.912	0.908

Table 2: Coverage probability of confidence intervals for $\theta_0(t),\ t = 1.1$

CR	n_1	$\alpha = 0.05$		$\alpha = 0.10$	
		EL	NA	EL	NA
0.10	30	0.927	0.872	0.877	0.863
	50	0.944	0.902	0.885	0.879
	80	0.947	0.946	0.897	0.890
	100	0.950	0.949	0.901	0.896
0.30	30	0.865	0.863	0.861	0.853
	50	0.936	0.934	0.880	0.875
	80	0.945	0.941	0.890	0.889
	100	0.952	0.946	0.899	0.897

Table 3: Coverage probability of simultaneous confidence bands for survival ratio under the Cox model on $[0.1, 4.0]$

CR	n_1	$z_0 = -1$	$z_0 = 0$	$z_0 = 1$
0.25	50	0.952	0.948	0.935
	100	0.956	0.950	0.930
0.50	50	0.950	0.946	0.930
	100	0.952	0.948	0.932
0.75	50	0.955	0.949	0.937
	100	0.951	0.946	0.934

Table 4: Coverage probability of simultaneous confidence bands for survival odds ratio under the Cox model on $[0.1, 4.0]$

CR	n_1	$z_0 = -1$	$z_0 = 0$	$z_0 = 1$
0.25	50	0.962	0.958	0.945
	100	0.966	0.960	0.950
0.50	50	0.965	0.966	0.940
	100	0.962	0.968	0.952
0.75	50	0.965	0.969	0.947
	100	0.961	0.966	0.954

Table 5: Coverage probability of simultaneous confidence bands for survival ratio under the additive risk model on $[0.2, 3.7]$

CR	n_1	z_0	EP	HW
0.25	50	0	0.948	0.955
	100	1	0.950	0.963
0.50	50	0	0.946	0.953
	100	1	0.948	0.952
0.75	50	0	0.949	0.947
	100	1	0.946	0.944

Table 6: Coverage probability of simultaneous confidence bands for survival odds ratio the additive risk model on $[0.3, 3.5]$

CR	n_1	$z_0 = -1$	$z_0 = 0$	$z_0 = 1$
0.25	50	0.962	0.945	0.950
	100	0.966	0.950	0.946
0.50	50	0.965	0.956	0.940
	100	0.962	0.952	0.948
0.75	50	0.965	0.947	0.939
	100	0.961	0.946	0.954

Appendix C

FIGURES

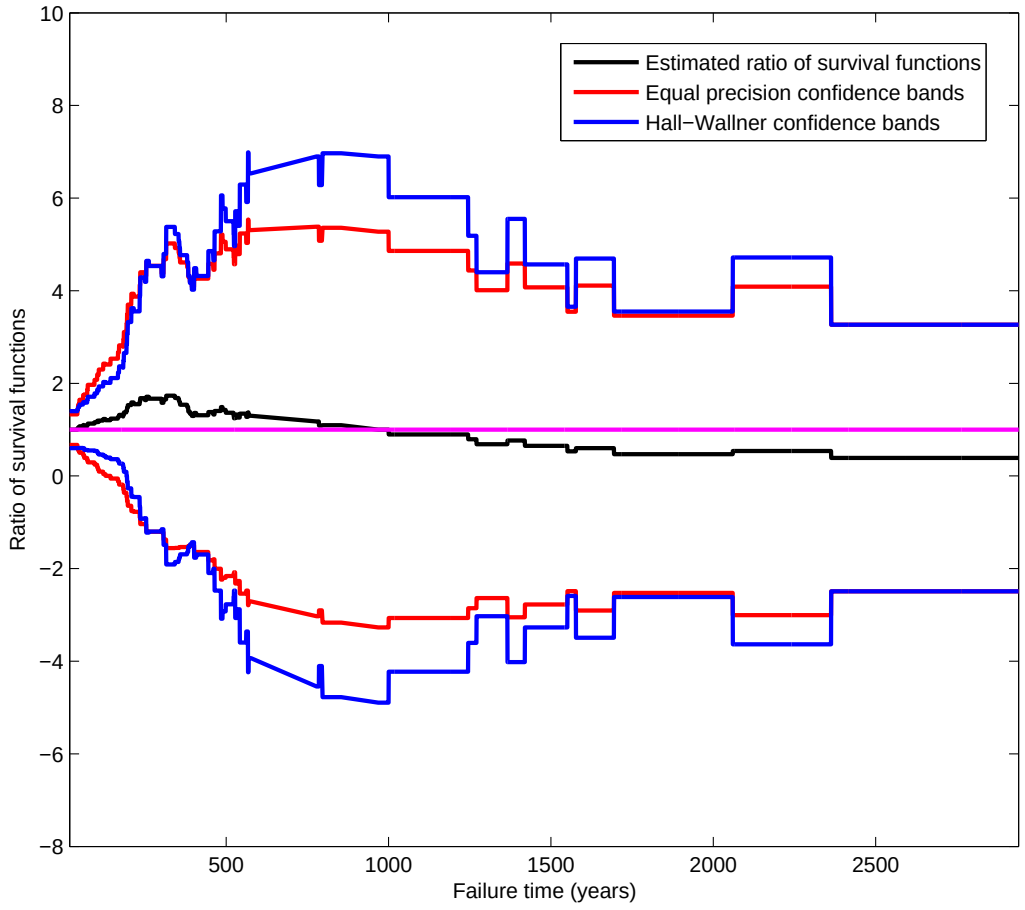

Figure 1: 95% simultaneous confidence bands for the survival ratio between the treatment group and the placebo group with the Mayo PBC data under the Cox model

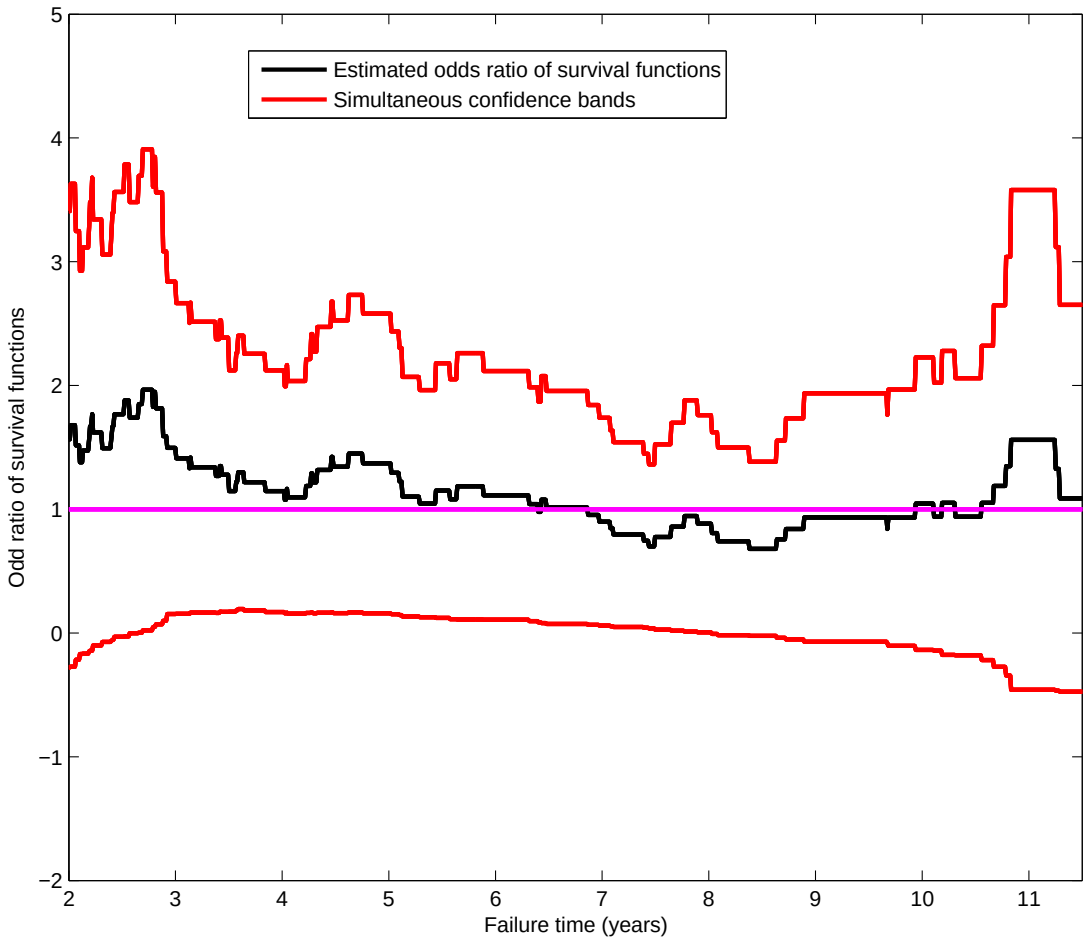

Figure 2: 95% simultaneous confidence bands for the survival odds ratio between the treatment group and the placebo group with the Mayo PBC data under the Cox model

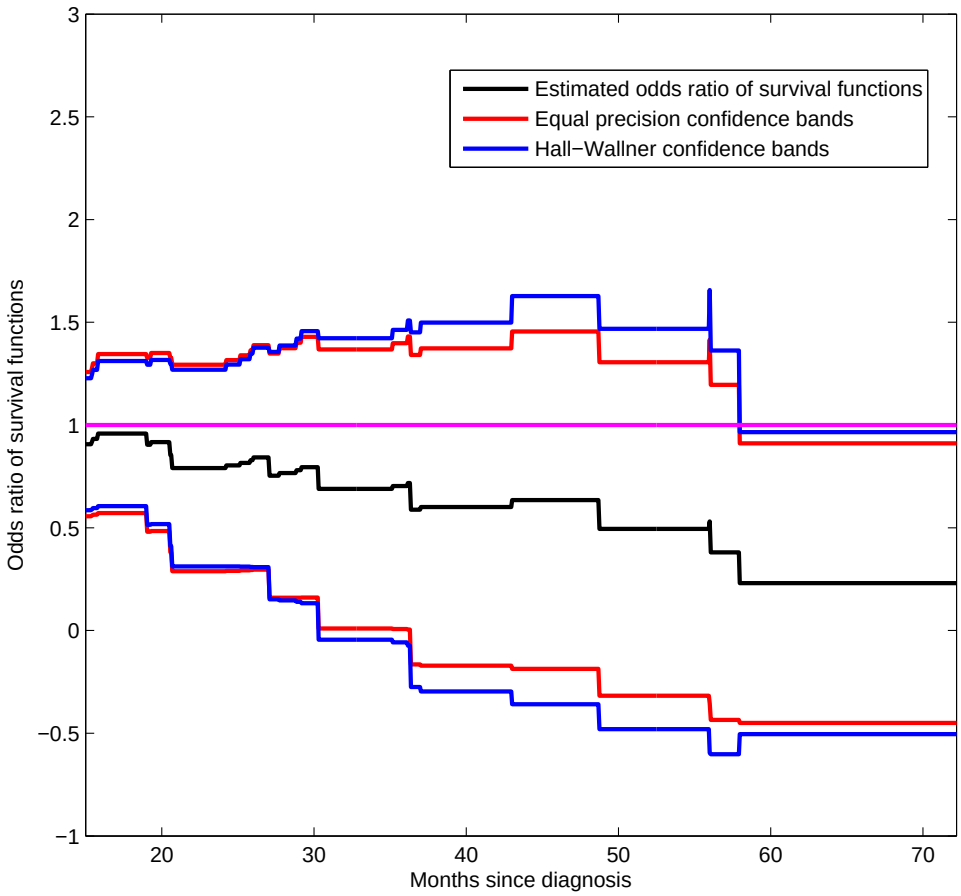

Figure 3: 95% simultaneous confidence bands for the survival ratio for patients with large spleen size using the CML data under the Cox model

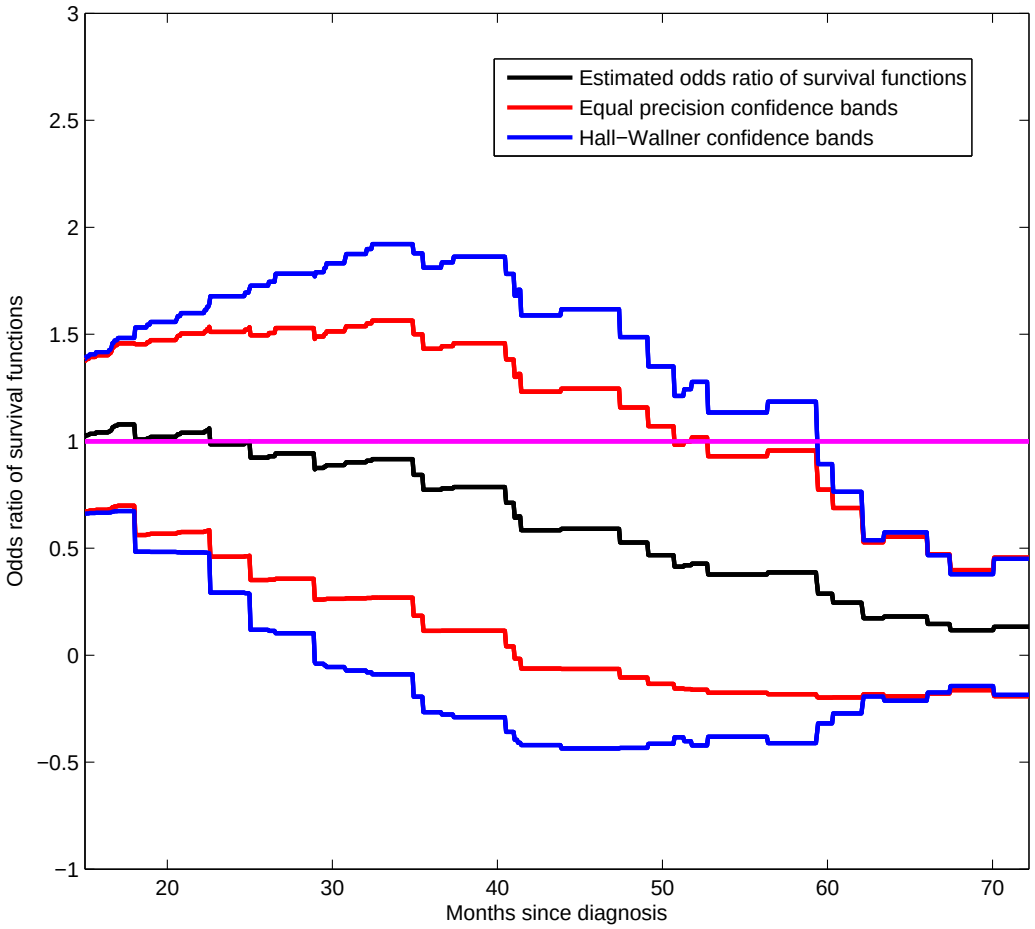

Figure 4: 95% simultaneous confidence bands for the survival ratio for patients with small spleen size using the CML data under the Cox model

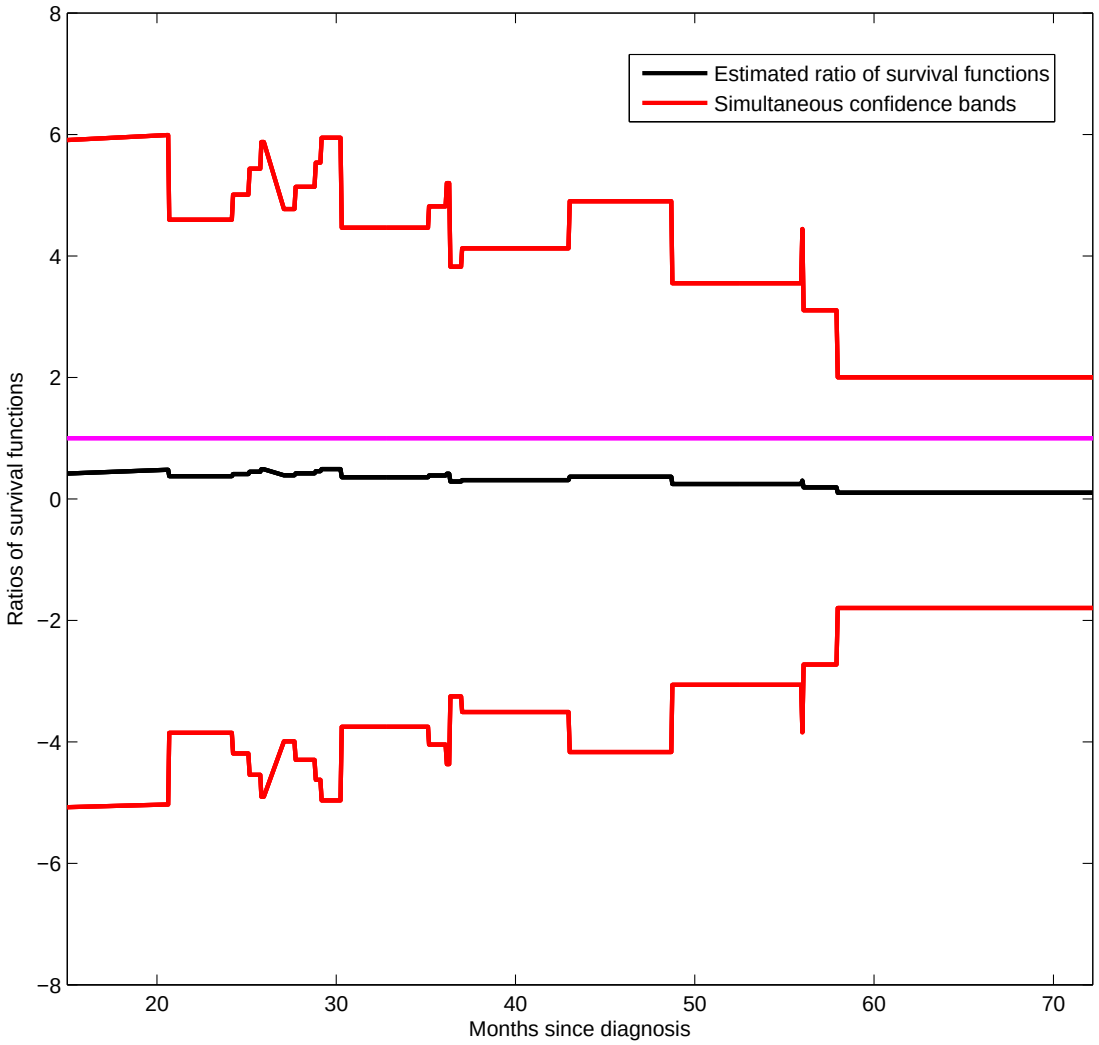

Figure 5: 95% simultaneous confidence bands for the survival odds ratio for patients with large spleen size using the CML data under the Cox model

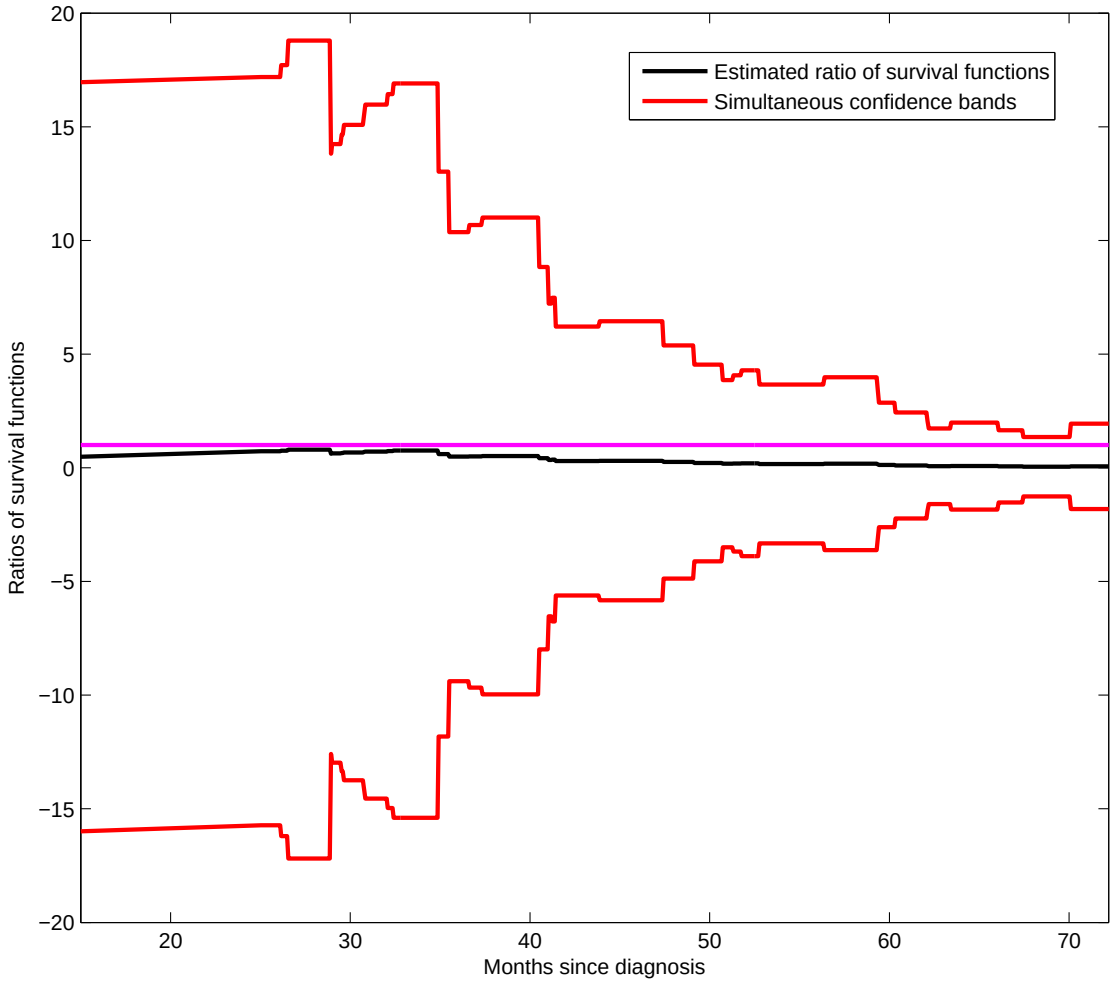

Figure 6: 95% simultaneous confidence bands for the survival odds ratio for patients with small spleen size using the CML data under the Cox model

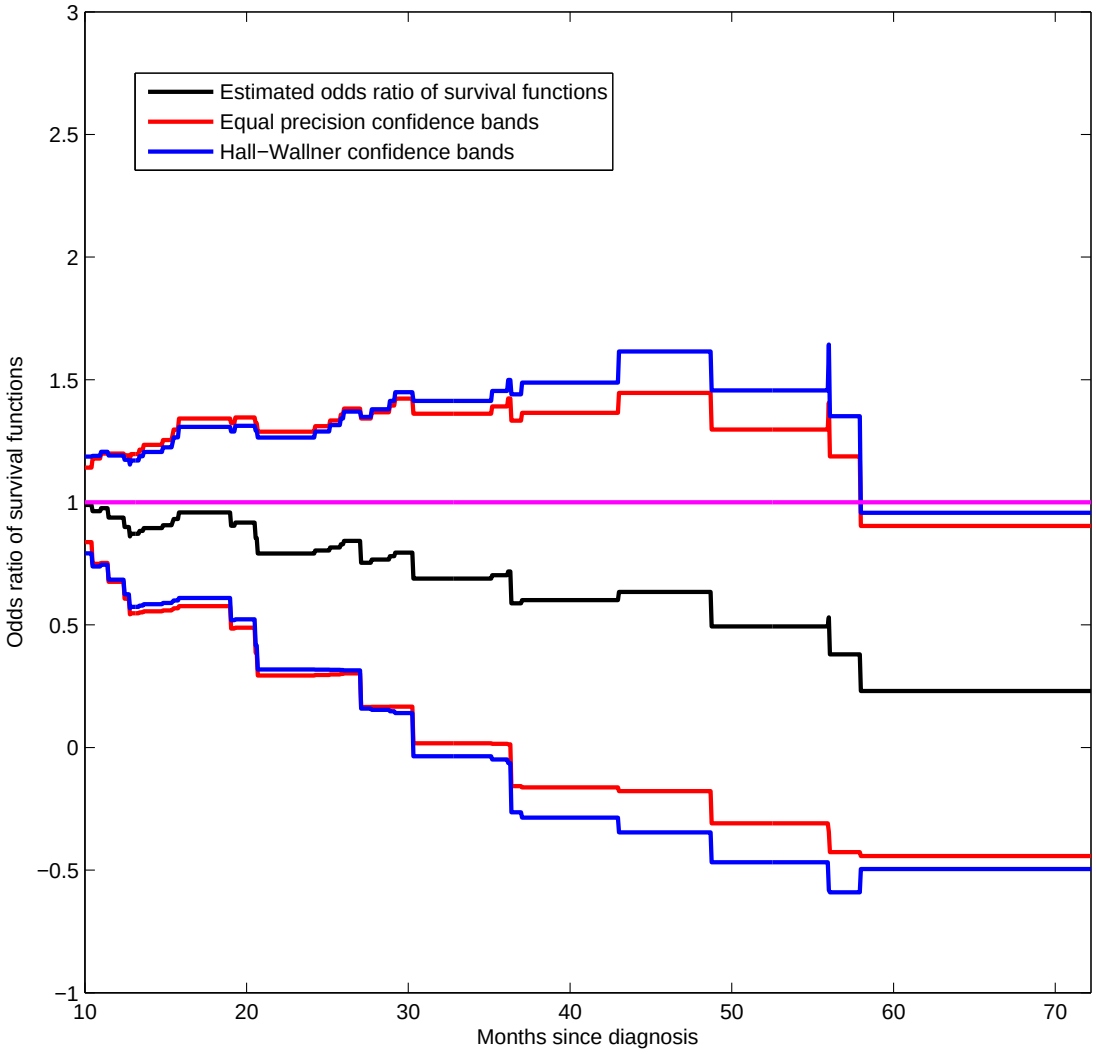

Figure 7: 95% simultaneous confidence bands for the survival ratio for patients with large spleen size using the CML data under the additive risk model

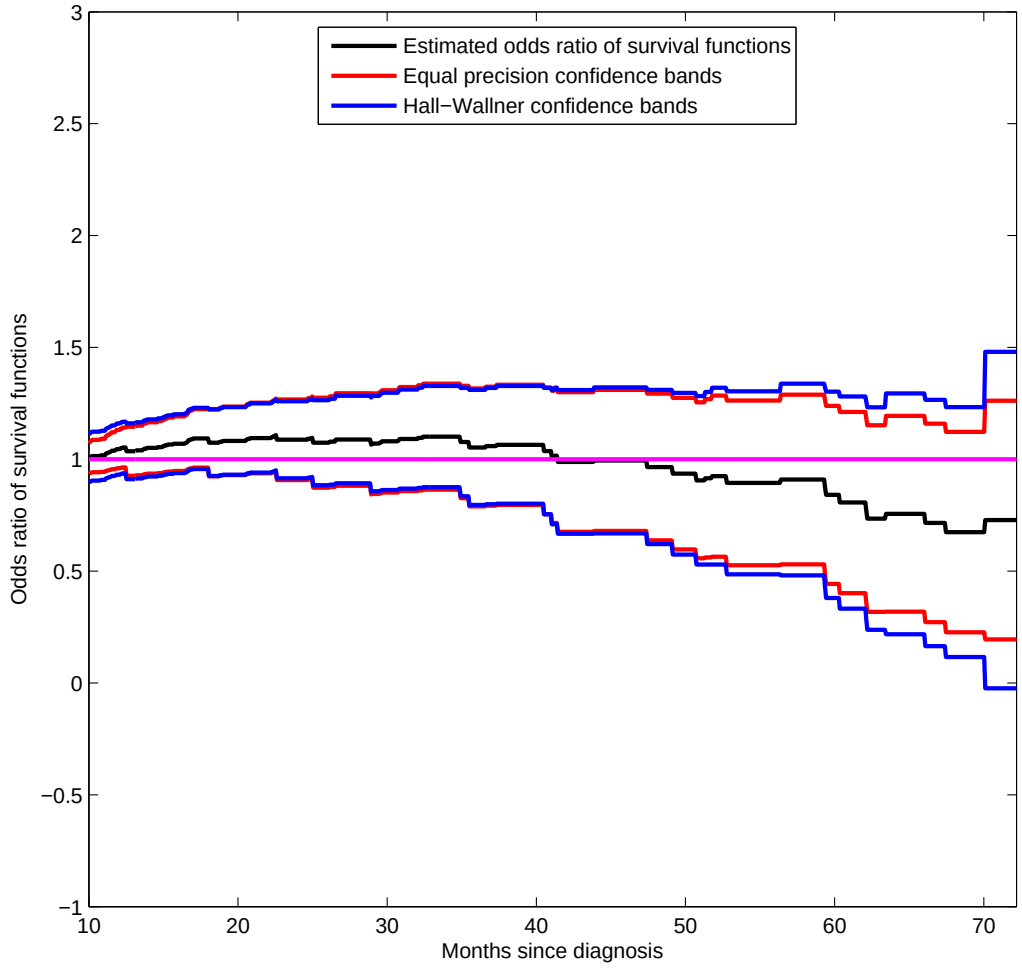

Figure 8: 95% simultaneous confidence bands for the survival ratio for patients with small spleen size using the CML data under the additive risk model

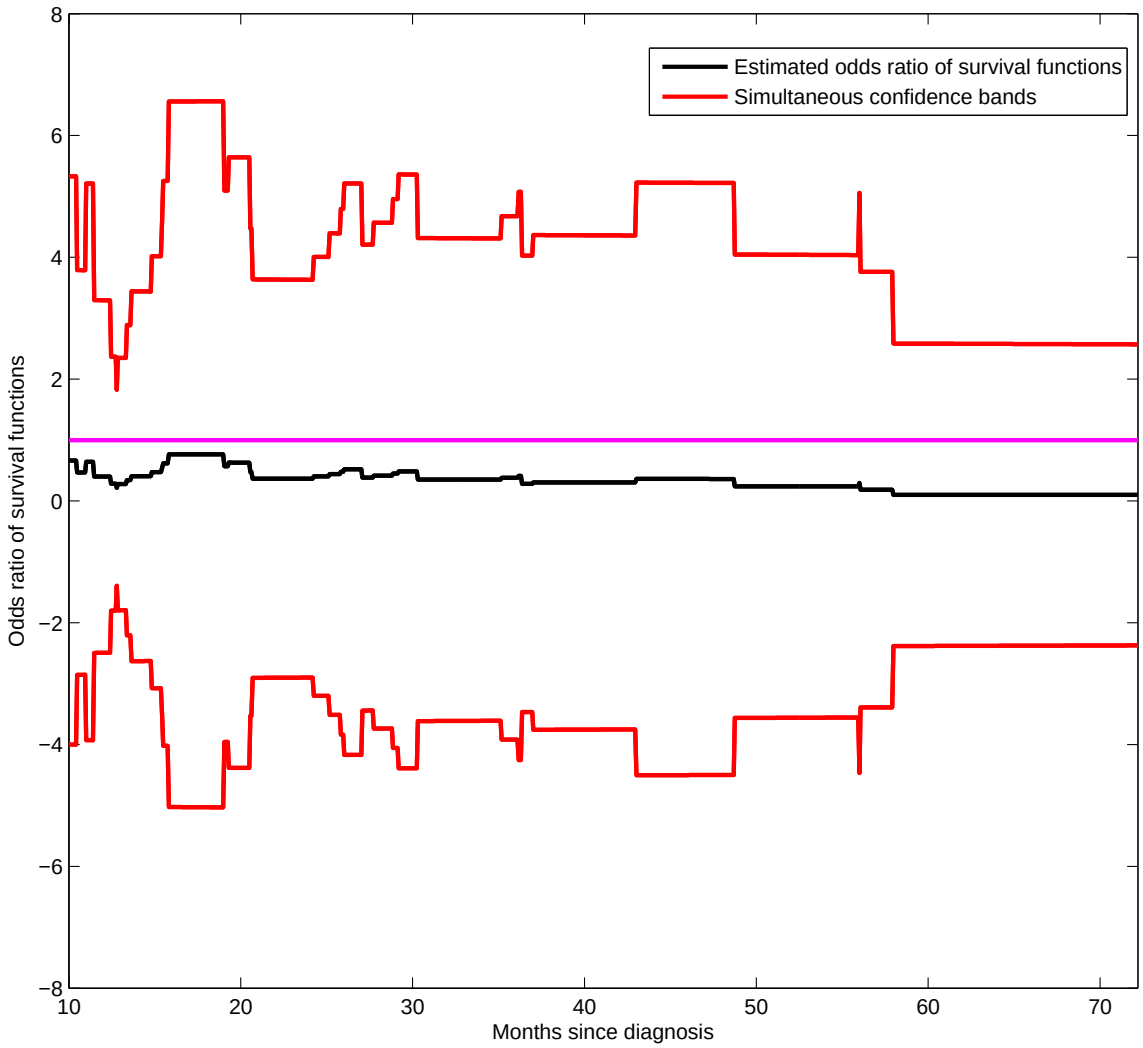

Figure 9: 95% simultaneous confidence bands for the survival odds ratio for patients with large spleen size using the CML data under the additive risk model

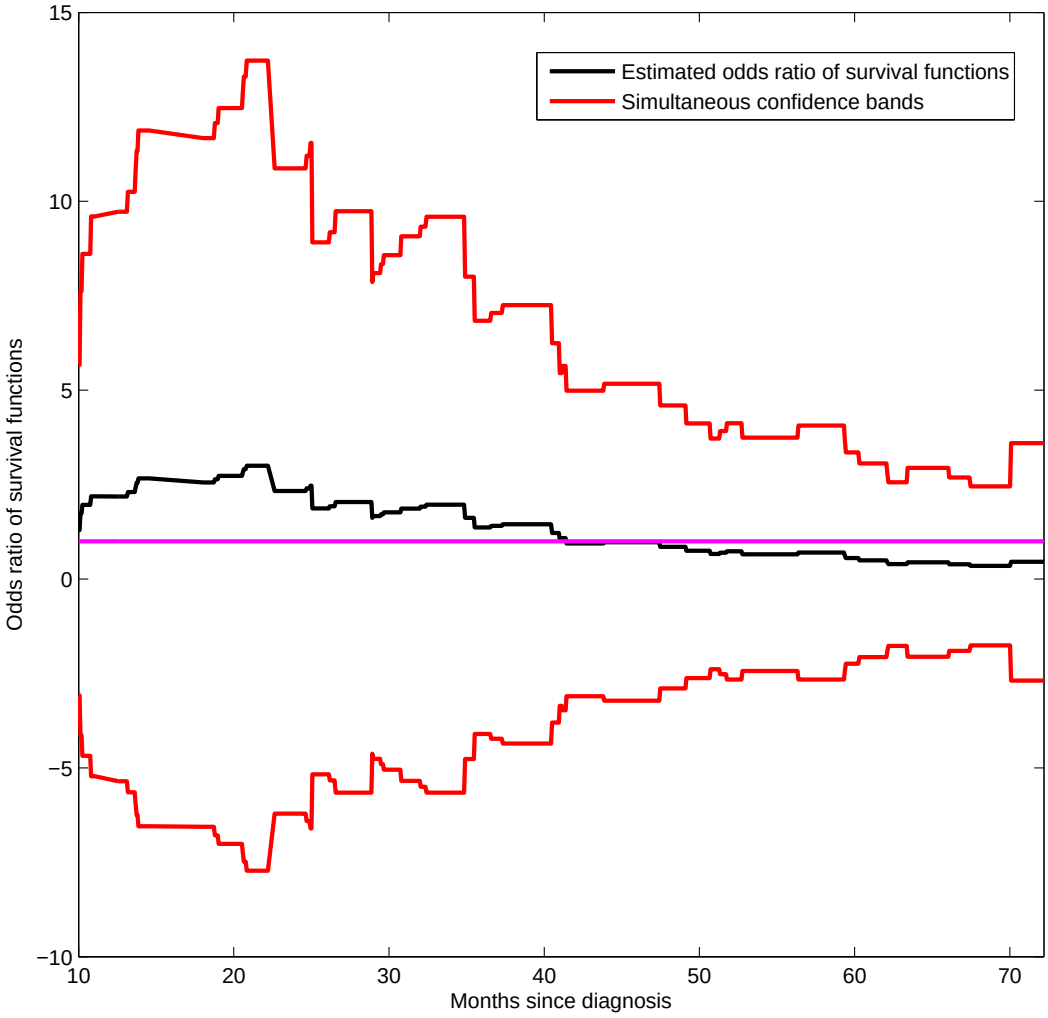

Figure 10: 95% simultaneous confidence bands for the survival odds ratio for patients with small spleen size using the CML data under the additive risk model

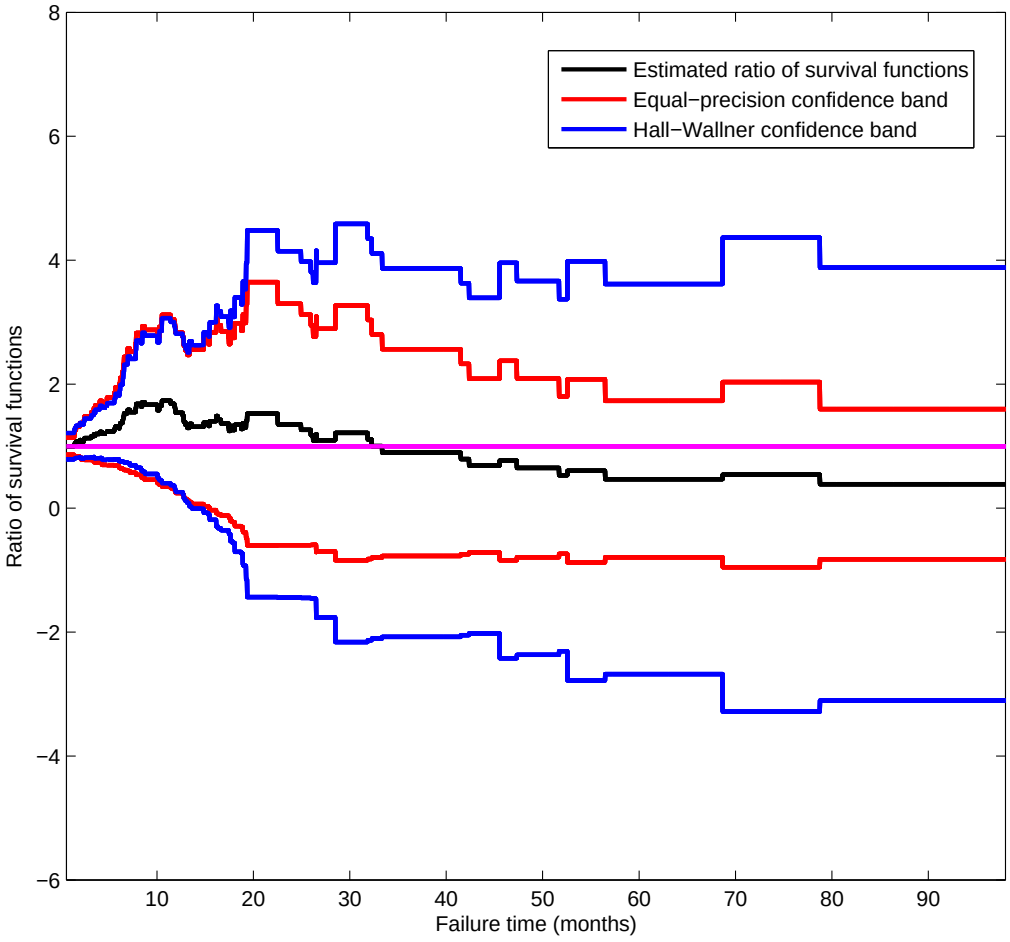

Figure 11: 95% simultaneous confidence bands for the survival ratio for patients with gastric cancer under the additive risk model

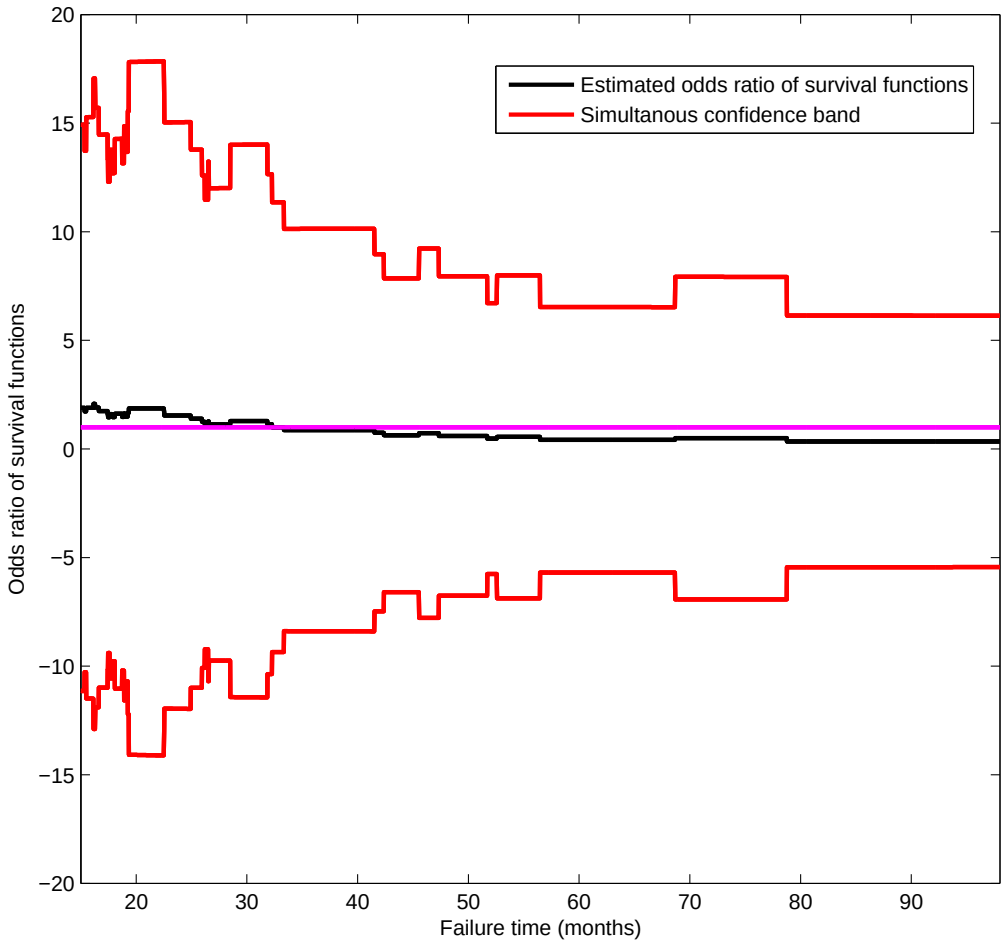

Figure 12: 95% simultaneous confidence bands for the survival odds ratio for patients with gastric cancer under the additive risk model

REFERENCES

[1] Aalen, O. O., A Model for Nonparametric Regression Analysis of Counting Processes. Lecture Notes in Statistics 2. *Springer:New York*, 1980.

[2] Aalen, O. O., A linear regression model for the analysis of life times. *Statistics in Medicine* Vol. 8, pp. 907–925, 1989.

[3] Andersen, P. K., Testing goodness of fit of Cox's regression and life mode. *Biometrics* Vol. 38, pp. 67–77, 1982.

[4] Andersen, P. K. and Gill, R. D., Coxs regression model for counting processes: A large-sample study. *Annals of Statistics*, Vol. 10, pp. 1100–1120, 1982.

[5] Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N., Statistical Models Based on Couting Prosses, *Springer, New York*, 1993.

[6] Arjas, E. A., Graphical method for assessing goodness of fit in Cox's proportional hazards model. *Journal of the American Statistical Association*, Vol. 83, pp. 204–212, 1988.

[7] Benichou, J., A review of adjusted estimators of attributeable risk. *Statistical Methods in Medical Research*, Vol. 10, pp. 195–216, 2001.

[8] Bie, O., Borgan, Ø. and Liestøl, K., Confidence interval and confidence bands for the cumulative hazard rate function and their small sample properties. *Scandinavian Journal of Statistics*, Vo. 14, pp. 221–233, 1987.

[9] Billingsley, P., Convergence of Probability Measures. 2nd ed. *Wiley: New York*, 1999.

[10] Breslow, N., Analysis of survival data under the proportional hazards model. *International Statistical Review*, Vol. 43, pp. 45–58, 1975.

[11] Breslow, N. E. and Day, N. E., Statistical Methods in Cancer Research 1, The Design and Analysis of Case-Control Studies. *IARC: Lyon*, 1980.

[12] Breslow, N. E. and Day, N. E., Statistical Methods in Cancer Research 2, The Design and Analysis of Case-Control Studies. *IARC: Lyon*, 1987.

[13] Bruzzi, P., Green, S. B., Byar, D. P., Brinton, L. A. and Schairer, C., Estimating the population attributable risk for multiple risk factors using case-control data. *American Journal of Epidemiology*, Vol. 122, pp. 904–914, 1985.

[14] Chen, K. and Ying, Z., A counterexample to a conjecture concerning the Hall-CWellner band. *Annals of Statistics*, Vol. 24, pp. 641–646, 1996.

[15] Chen, L., Lin, D. Y., and Zeng, D., Attibutable fraction functions for censored event times. *Biometrika*, Vol. 97, pp. 713–726, 2010.

[16] Chen, Y. Q., Hu, C. and Wang, Y., Attributable risk functions in the proportional hazards model for censored time-to-event. *Biostatistics*, Vol. 7, pp. 515–529, 2006.

[17] Cox, D. R., Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, Vol. 34, pp. 187–220, 1972.

[18] Cox, D. R., Partial likelihood. *Biometrika*, Vol. 62, pp. 269–276, 1975.

[19] Csörgö, S. and Horváth, P., The rate of strong uniform consistency for the product-limit estimator. *Probability Theory and Related Fields*, Vol. 62, pp. 411-426, 1983.

[20] Cummings, P., Early Exposure to Marijuana and Risk of Later Drug Use. *Journal of the American Medical Association*, Vol. 290(3), pp. 329, 2003.

[21] Dabrowska, D. M., Doksum, K. A. and Song, J-K., Graphical comparison of cumulative hazards for two populations. *Biometrika*, Vol. 76, pp. 763–773, 1989.

[22] Dabrowska, D. M., K. A. Doksum, Feduska, N. J., Husing, R. and Neville, P., Methods for comparing cumulative hazard functions in a semi-proportional hazard model. *Statistics in Medicine*, Vol. 11, pp. 1465–1476, 1992.

[23] Dauxois, J.Y. and Kirmani, S.N.U.A., Testing the proportional odds model under random censoring. *Biometrika*, Vol. 90, pp. 913–922, 2003.

[24] Einmahl, J.H. and McKeague, I.W., Confidence tubes for multiple quantile plots via empirical likelihood. *Annals of Statistics*, Vol. 27, pp. 1348–1367, 1999.

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

[26] Gale, R. B., Helmann, R., Zhang, M. J., Hasford, J., Goldman, J., Heimpel, H., Klein, J. P., Kolb, H. J., McGlave, P. B., Passweg, J. R., Rowlings, P. A., Sobocinski, K. A., Horowitz, M. M. and the German CML Study Group., Survival with bone marrow transplantation versus hydroxyurea or interferon for chronic myelogenous leukemia (CML) *Blood*, Vol. 91, pp. 1810–1819, 1998.

[27] Gastrointestinal Tumor Study Group: Schein, P. D., Bruckner, H. W., Dougluass, H. O., Mayer, R. et al., A comparison of combination chemotherapy and combined modality therapy for locally advanced gastric carcinoma. *Cancer*, Vol. 49, pp. 1771-1777, 1982.

[28] Graubard, B. I. and Fears, T. R., Standard errors for attibutable risk for simple and complex sample designs. *Biometrics*, Vol. 61, pp. 847-855, 2005.

[29] Greenland, S., Estimation of population attributable fractions from fitted incidence ratios and exposure survey data, with an application to electromagnetic fields and childhood leukemia. *Biometrics*, Vol. 57, pp. 182–188, 2001.

[30] Hall, P. and Owen, A.B., Empirical likelihood confidence bands in density estimation. *Journal of Computational and Graphical Statistics*, Vol. 18, pp. 121-140, 1993.

[31] Hall, W. J. and Wellner, J. A., Confidence bands for a survival curve from censored data. *Biometrika*, Vol. 67, pp. 133–143, 1980.

[32] Hassan, E., Recall bias can be a threat to retrospective and prospective research designs. *The Internet Journal of Epidemiology*, Vol. 3 (2), 2006.

[33] Hollander, M., McKeague, I.W., and Yang, J., Likelihood ratio-based confidence bands for survival functions. *Journal of the American Statistical Association*, Vol. 18, pp. 121-140, 1997.

[34] Huffer, F. W. and McKeague, I. W., Weighted least squares estimation for Aalen??s additive risk model. *Journal of the American Statistical Association*, Vol. 86, pp. 38–53, 1991.

[35] Jing, B.Y., Two sample empirical method., *Statistics and Probability Letters*, Vol. 25, pp. 95-104, 1995.

[36] Kalbfleisch, J.D. and Prentice, R.L., Estimation of the average hazard ratio. *Biometrika*, Vol. 68, pp. 105–112, 1981.

[37] Kalbfleisch, J. D. and Prentice, R. L., The statistical analysis of failure time data. 2nd ed. *Wiley: New York*, 2002.

[38] Kaplan, E. L. and Meier, P., Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, Vol. 53, pp. 457–481, 1958.

[39] Kim, J. and Lee, S. Y., Two-sample goodness-of-fit tests for additive risk models with censored observations. *Biometrika*, Vol. 85, pp. 593–603, 1998.

[40] Kulich, M. and Lin, D. Y., Additive hazard regression with covariate measurement error. *Journal of the American Statistical Association*, Vol. 95, pp. 238–248, 2001.

[41] Lai, T. L. and Ying. Z., Estimating a distriubtion function with truncated and censored data. *The Annuals of Statistics*, Vol. 19, pp. 417–442, 1991.

[42] Lee, J. and Hyun, S., Confidence bands for the difference of two survival functions under the additive risk model. *Journal of Applied Statistics*, Vol. 38, pp. 785–797, 2011.

[43] Levin, M. L., The occurrence of lung cancer in man. *Acta Unio Internationalis Contra Cancrum*, Vol. 9, pp. 531–541, 1953.

[44] Li, G., On nonparametric likelihood ratio estimation of survival probabilities for censored data. *Statistics and Probability Letters*, Vol. 25, pp. 95-104, 1995.

[45] Li, G., Li, R. and Zhou, M., Empirical Likelihood in Survival Analysis. Contemporary Multivariate Analysis and Design of Experiments J. Fan and G. Li (eds). *World Scientific: Singapore*, 2005.

[46] Li, G. and Van Keilegom, I., Likelihood ratio confidence bands in non-parametric regression with censored data. *Scandinavian Journal of Statistics*, Vol. 29, pp. 547–562, 2002.

[47] Lin, D. Y., Fleming, T. R. and Wei, L. J., Confidence bands for survival curves under the proportional hazards model. *Biometrika*, Vol. 81, pp. 73–81, 1994.

[48] Lin, D. Y., Oakes, D. and Ying, Z., Additive hazards regression for current status data. *Biometrika*, Vol. 85, pp. 289–298, 1998.

[49] Lin, D. Y., Wei, L. J. and Ying, Z., Checking the Cox model with cumulative sum of Martingale-based residuals. *Biometrika*, Vol. 80, pp. 557–572, 1993.

[50] Lin, D.Y., Wei, L.J., and Ying, Z., Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika*, Vol. 80, pp. 557–572, 1999.

[51] Lin, D. Y. and Ying, Z., Semiparametric analysis of the additive risk model. *Biometrika*, Vol. 81, pp. 61–71, 1994.

[52] Lin, D. Y. and Ying, Z., Additive hazards regression models for survival data. In: Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis. *Springer: New York*, 1997.

[53] Lo, S.H., Mack, Y.P. and Wang, J.L., Density and Hazard rate estimation for censored data via strong representation of the Kaplan–Meier estimator. *Probability Theory and Related Fields*, Vol. 80, pp. 461–473, 1999.

[54] McKeague, I. W. and Sasieni, P. D., A partly parametric additive risk model. *Biometrika*, Vol. 81, pp. 501–514, 1994.

[55] McKeague, I.W. and Zhao, Y., Simultaneous confidence bands for ratios of survival functions via empirical likelihood. *Statistics and Probability Letters*, Vol. 60, pp. 405-415, 2002.

[56] McKeague, I.W. and Zhao, Y., Comparing distribution functions via empirical likelihood. *International Journal of Biostatistics*, article 5, pp. 1-20, 2005.

[57] Murphy, S., Likelihood ratio-based confidences in survival analysis. survival functions. *Journal of the American Statistical Association*, Vol. 90, pp. 1399-1405, 1995.

[58] Nair, V. N., Confidence bands for survival functions with censored data: A comparative study. *Technometrics*, Vol. 26, pp. 265–275, 1984.

[59] Owen, A. B., Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, Vol. 75, pp.237-249, 1988.

[60] Owen, A. B., Empirical likelihood and confidence regions. *Annals of Statistics*, Vol. 18, pp. 90-120, 1990.

[61] Owen, A. B., Empirical Likelihood. *Chapman and Hall/CRC: Boca Raton*, 2001.

[62] Parzen, M. I., Wei, L. J. and Ying, Z., Simultaneous confidence intervals for the difference of two survival functions. *Scandinavian Journal of Statistics*, Vol 24, 309–314, 1997.

[63] Qin, J., Semi-empirical likelihood ratio confidence intervals for the difference of two sample means. *Statistics and Probability Letters*, Vol. 46, pp. 117-126, 1994.

[64] Qin, Y.S., Semi-empirical likelihood ratio confidence intervals for various differences of two populations. *Annals of the Institute of Statistical Mathematics*, Vol. 33, pp. 135-143, 1997.

[65] Qin, Y.S. and Zhao, L.C., Empirical likelihood ratio confidence intervals for various differences of two populations. *Journal of Systems Science and Mathematical Science*, Vol. 13, pp. 23-30, 2000 (in Chinese).

[66] Ren, J.J., Weighted empirical likelihood in some two-sample semiparametric models with various types of censored data. *Annals of Statistics*, Vol. 36, pp. 145–166, 2008.

[67] Rothman, K. J. and Greenland, S., Modern Epidemiology, 2nd ed. *Lippincott-Raven: Philadelphia*, 1998.

[68] Scheike, T. H., The additive nonparametric and semiparametric Aalen model as the rate function for a counting process. *Lifetime Data Analysis*, Vol. 8, pp. 247–262, 2002.

[69] Schemper, M., Cox analysis of survival data with nonproportional hazard functions. *The Statistician*, Vol. 41, pp. 455–465, 1992.

[70] Sen, P. K. and Singer, J. M., Large Sample Methods in Statistics: An Introduction with Applications. *Chapman: New York*, 1993.

[71] Shen, J. and He, S., Empirical likelihood for the difference of two survival functions under right censorship. *Statistics and Probability Letters*, Vol. 76, pp. 169-181, 2006.

[72] Shen, Y. and Cheng, S. C., Confidence bands for cumulative incidence curves under the additive risk model. *Biometrics*, Vol. 55, pp. 1093–1100, 1999.

[73] Silverberg, M. J., Smith, M. W., Chmiel, J. S., Detels, R., Margolick, J. B., Rinaldo, C. R., O′Brien, S. J. and Muñoz, A. Fraction of cases of acquired immunodeficiency syndrome prevneted by the interactions of identified restriction gene variations. *American Journal of Epidemiology*, Vol. 159, pp. 232–241, 2004.

[74] Song, M.U., Jeong, D.M. and Song, J.K. Confidence bands for survival curve under the additive risk model. *Journal of the Korean Statistical Society*, Vol. 26(4), pp. 429–443, 1996.

[75] Thomas, D.R. and Grunkemeier, G.L., Confidence interval estimation for survival probabilities for censored data. *Journal of the American Statistical Association*, Vol. 70, pp. 865-871, 1975.

[76] Wei, G. and Schaubel, D. E., Estimating cumulative treatment effects in the presence of nonproportional hazards. *Biometrics*, Vol. 64, pp. 724–732, 2008.

[77] Whittemore, A. S., Statistical methods for estimating attributable risk from retrospective data. *Statistics in Medicine*, Vol. 1, pp. 229–243, 1982.

[78] Xu, R. and O′Quigley, J., Estimating average regression effect under nonproportional hazards. *Biostatistics*, Vol. 1, pp. 423–439, 2000.

[79] Yang, S. and Prentice, R. L., Semiparametric inference in the proportional odds regression model. *Journal of the American Statistical Association*, Vol. 94, pp. 125–136, 1999.

[80] Yang, S. and Prentice, R. L., Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. *Biometrika*, Vol. 92, pp. 1–17, 2005.

[81] Yin, G. and Hu, J., Two Simulation Methods for Constructing Confidence Bands Under the Additive Risk Model. *Journal of Biopharmaceutical Statistics*, Vol. 14(2), pp. 389–402, 2004.

[82] Ying, Z., Jung, S. H. and Wei, L. J., Survival analysis with median regression models. *Journal of the American Statistical Association*, Vol. 90, pp. 178–184, 1995.

[83] Yip, P. S. F., Zhou, Y., Lin, D. Y. and Fang, X. Z., Estimation of population size based on additive hazards models for continuous-time recapture experiments. *Biometrics*, Vol. 55, pp. 904–908, 1999.

[84] Yuen, K. C. and Burke, M. D., A test of fit for a semiparametric additive risk model. *Biometrika*, Vol. 84, pp. 631–639, 1997.

[85] Zhang, J and Yu, K.F., What's the Relative Risk?: A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. *Journal of the American Statistical Association*, Vol. 280(19), pp. 1690-1691, 1998.

[86] Zhang, M.J. and Klein, J., Confidence bands for the difference of two survival curves under proportional hazards model. *Lifetime Data Analysis*, Vol. 7, pp. 234-254, 2001.