國立臺灣大學理學院數學研究所

博士論文

Department of Mathematics

College of Science

National Taiwan University

Doctoral Dissertation

時間相關 AUC 之方法論及其應用

Methodology for the Time-Dependent AUC and Its Applications

洪 弘

Hung Hung

指導教授: 江金倉 博士

Advisor: Chin-Tsang Chiang, Ph.D.

中華民國九十八年四月

April, 2009

Table of Contents

Table of Contents iii

List of Tables v

List of Figures viii

Acknowledgements ix

摘要 x

Abstract xii

1 Introduction 1
- 1.1 The ROC Curve Analysis . . . 1
- 1.2 Motivating Examples . . . 3
  - 1.2.1 Angiography Coronary Artery Disease (CAD) Study . . . 3
  - 1.2.2 AIDS Clinical Trials Group (ACTG) 175 Study . . . 4
- 1.3 Time-Dependent Monitoring Rule . . . 4
- 1.4 Existing Methods for Time-Dependent Vital Status . . . 5
- 1.5 Main Contributions . . . 7

2 Time-Dependent AUC and PAUC 10
- 2.1 Nonparametric Estimation . . . 11
- 2.2 Asymptotic Properties . . . 13
- 2.3 Inference Procedures . . . 18

3 Semiparametric Regression Model for the Time-Dependent AUC 21
- 3.1 Estimation . . . 22
- 3.2 Asymptotic Properties . . . 23
- 3.3 Inference Procedures . . . 25

4 Optimal Composite Markers **28**
4.1 Estimation Procedures 29
4.2 Asymptotic Properties 31

5 Simulations and Applications **41**
5.1 Monte Carlo Simulations 41
5.1.1 Scenario I - Time-Dependent AUC and PAUC 41
5.1.2 Scenario II - Time-Dependent AUC Regression Model 43
5.1.3 Scenario III - Optimal Composite Markers 45
5.2 Applications 46
5.2.1 ACTG 175 Study 46
5.2.2 Angiography CAD Study 49

6 Discussion **74**
6.1 Concluding Remarks 74
6.2 Future Works 76
6.2.1 Multiple Biomarker Comparison 76
6.2.2 Optimality in Classification and Prediction 78
6.2.3 Partly Conditional Time-Dependent AUC 79

Bibliography **80**

List of Tables

5.1 The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under totally independent censoring 51

5.2 The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under totally independent censoring 52

5.3 The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under totally independent censoring 53

5.4 The empirical coverage probabilities of 0.95 simultaneous confidence bands under totally independent censoring 54

5.5 The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under marker-dependent censoring 55

5.6 The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under marker-dependent censoring 56

5.7 The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under marker-dependent censoring 57
5.8 The empirical coverage probabilities of 0.95 simultaneous confidence bands under marker-dependent censoring 58
5.9 The averages (Mean) and standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) of 0.95 pointwise confidence intervals of $\widehat{\gamma}_t$ 59
5.10 The averages (Mean) and standard deviations (SD) of 500 estimates $\widehat{\theta}_t^{(T)}(z_0, z_0)$'s, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) of 0.95 pointwise confidence intervals 60
5.11 The averages (Mean) and standard deviations (SD) of 500 estimates $\widehat{\theta}_t^{(W)}(z_0, z_0)$'s, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) of 0.95 pointwise confidence intervals 61
5.12 The empirical coverage probabilities of 0.95 simultaneous confidence bands with the sample sizes of 250 and 500, and the censoring rates of 30% and 50% 62
5.13 The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under totally independent censoring with $c.r. = 30\%$ 63
5.14 The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under totally independent censoring with $c.r. = 50\%$ 64
5.15 The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under marker-dependent censoring with $c.r. = 30\%$ 65

5.16 The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under marker-dependent censoring with $c.r. = 50\%$ 66
5.17 The estimates of the coefficients in the linear predictor tHcy+β_{t2}CRP+ β_{t3}SAA + β_{t4}IL-6 (95% bootstrap confidence intervals) at the selected time points 72
5.18 Estimates of $\theta_t^{(\text{CRP})}$, $\theta_t^{(\text{SAA})}$, $\theta_t^{(\text{IL-6})}$, $\theta_t^{(\text{tHcy})}$, and $\theta_t(\beta_t)$ at the selected time points 73

List of Figures

5.1 The estimated time-dependent PAUCs (solid curve) with the corresponding 0.95 pointwise (dotted curve) and simultaneous confidence bands (dashed curve) 67
5.2 The estimated curves for the difference of the time-dependent PAUCs between non-therapy and therapy patients (solid curve) with the corresponding 0.95 pointwise (dotted curve) and simultaneous confidence bands (dashed curve) 68
5.3 The estimated coefficient and time-dependent AUC curves (solid curve) with the corresponding 0.95 pointwise (dotted curve) and simultaneous confidence bands (dashed curve) 69
5.4 The estimated time-dependent AUCs (solid curve) under totally independent censoring with the corresponding 0.95 pointwise (dotted curve) and simultaneous (dashed curve) confidence bands 70
5.5 The estimated time-dependent AUCs (solid curve) under marker-dependent censoring with the corresponding 0.95 pointwise (dotted curve) and simultaneous (dashed curve) confidence bands 71

Acknowledgements

I wish to express my sincere appreciation to my parents, teachers, and friends.

摘要

ROC (receiver operating characteristic) 曲線分析已經被廣泛地應用於早期疾病診斷之評估。其中 AUC (area under the ROC curve) 以及 PAUC (partial area under the ROC curve) 是兩個最常使用的指標。近來在實務應用上, 隨時間變化之疾病狀態 (disease status) 的資料已經越來越常見。不同於以往單純的二分類 (有病或沒病), 時間相關之疾病狀態是由一個特定事件的發生時間所定義出來的。給定任一個時間點, 資料被分爲在此時間點之前有發生疾病以及沒有發生疾病的兩個群體。有鑒於此, 傳統的 ROC 曲線分析也必須推廣到時間相關 ROC 曲線分析。由於事件發生的時間有可能被截切 (censored) 以至於無法知道確切的發生時間, 在使用時間相關 ROC 曲線分析時, 最大的挑戰來自於如何利用不完整的資料進行統計推論。另一個常遇到的問題是使用單一生物指標來分類疾病狀態往往無法達到預期的水準。在很多情況下, 受試者會被同時觀測到多重生物指標。如何結合多重生物指標以增進分類的能力也是一個很重要的議題。

不同於現有的方法, 我針對時間相關之 AUC 以及 PAUC 提出一系列的無母數估計方法。由於這些估計方法有明確的表達式, 因此不論是對計算效率還是大樣本理論之推導都有很重要的貢獻。此外, 我也提出了廣義線性模型以分析時間相關之 AUC 與其它變數之間的相關性。最後, 在承認條件存活分布函數滿足

EGLM (extended generalized linear model) 並且不需指定連結函數確切形式的情況下, 我在這本論文裡對於最佳合成指標也提出了兩個無母數估計方法。所有建立的推論方法都應用在 AIDS Clinical Trials Group (ACTG) 175以及 Angiography Coronary Artery Disease (CAD) 這兩筆臨床資料以驗證其實用性。

關鍵詞: AUC, 分類, 疾病狀態, 最佳合成指標, 預測, ROC, 存活時間

Abstract

To evaluate the performance of test results in early detection of disease, the receiver operating characteristic (ROC) curves are widely used. The area under the ROC curve (AUC) and the partial area under the ROC curve (PAUC) are the most popular summary measures for its generality and ease of probability interpretation. In applications, data with the binary time-varying disease status are frequently encountered. The cases and controls in the ROC analysis are more suitable defined over time. A major challenge in dealing this issue is that the failure status of some individuals might not be available due to censoring. To further increase classification ability of multiple biomarkers, research interests usually focus on seeking combinations of these biomarkers with the highest ROC curve.

In contrast to the existing methods, we propose nonparametric estimators for the time-dependent AUC and PAUC with explicit expressions and a rigorous theoretical development for these methods. Moreover, we use a generalized linear model with time-varying coefficients to characterize the time-dependent AUC as a function of covariate values.

For the parameter functions and the related classification accuracies, the estimation and inference procedures are also proposed. Under the validity of an extended generalized linear model (EGLM) with time-varying coefficients and an unknown link function for the conditional survival distribution, two nonparametric procedures are proposed to estimate the optimal composite markers based on the estimation procedures of the time-dependent AUC. Two empirical examples from the AIDS Clinical Trials Group (ACTG) 175 study and the Angiography Coronary Artery Disease (CAD) study are used to illustrate the usefulness of our methods. Finally, some concluding remarks and further research topics of interests are devoted in this thesis.

Key words and phrases: AUC, classification, disease-status, optimal composite biomarker, prediction, ROC, survival time

Chapter 1

Introduction

1.1 The ROC Curve Analysis

Decision-making is an important issue in many fields such as psychology, radiology, medicine, etc. For example, in clinical preventive medicine preoperative diagnostic tests are medically necessary and implemented to determine those patients for whom surgery is beneficial. For the sake of cost-saving or performance improvement, new diagnostic tests are often introduced and the classification accuracies of them are evaluated and compared with the existing ones. The decision is usually to classify a subject to be diseased or underwent operation if the test result Y is greater than a threshold value y. In practice, two conditional probabilities are applied to evaluate the performance of such classification criterion: the true positive rate (sensitivity) $TPR(y) = P(Y > y|D = 1)$ and the false positive rate (1-specificity) $FPR(y) = P(Y > y|D = 0)$ with D being the binary disease status where $D = 1$ indicates diseased and $D = 0$ otherwise. Although a good classifier should possess high $TPR(y)$ and low $FPR(y)$, it is observed that both measures are decreasing functions of y. The trade-off between $TPR(y)$ and $FPR(y)$ then become crucial in the assessment. The ROC curve, a graph of $(FPR(y), TPR(y))$ points, is widely used to display the relationship between $TPR(y)$ and $FPR(y)$. A merit to favor the ROC curve is that the choice of possible value y may depend on different research purposes and related

cost. Without cosidering a specific threshold value, the ROC curve shows the inherent classification capability of a biomarker. Interestingly, the test result Y and any strictly increasing transformation $g(Y)$ have the same ROC curve. This invariance characteristic of the ROC curve in measurement scale further provides a suitable base to compare different biomarkers. Geomerically, the more the curve moves toward the point $(0, 1)$ the better a biomarker performs, while a straight line pass through $(0, 0)$ and $(1, 1)$ indicates a useless one.

In applications, one of the most popular overall performance measures is the area under the ROC curve (AUC). It has the probability explanation that the considered marker of a randomly selected diseased case is greater than that of a non-diseased one, i.e., $P(Y_i > Y_j | D_i = 1, D_j = 0)$. Obviously, a perfect marker will have the AUC value 1 while a poor one takes a value close to 0.5. What is the value of the AUC should a good or acceptable marker possess? The standard is quite different and might depend on different research purposes. A biomarker with the AUC value greater than 0.75 is regarded as being indicative of disease in medical practice. In emergency medicine, however, a useful diagnostic test should have the AUC value at least 0.9. Since the AUC is an overall evaluation, relevant information might not be entirely captured in some cases. For example, two crossed ROC curves might have the same AUC but totally different performances. It is more reasonable to evaluate the performance of a biomarker over the practically acceptable portion of the ROC curve only. The partial AUC (PAUC) for FPR over the relevant interval was adopted by McClish (1989) as the summary measure. Jian, Metz, and Nishikawa (1996) argued that women with false-negative findings at mammography cannot be benefited from timely treatment of the cancer and, hence, suggested using the PAUC with restricted true positive range in their applied data. As mentioned by Dwyer (1997), the PAUC is a regional analysis of the ROC curve intermediate between the AUC and individual points on the ROC curve. In view of these points, it is more useful to consider the PAUC as alternative summary measure due to its sensitivity and flexibility. The PAUC with $FPR(y) < \alpha$, $\alpha \in (0, 1]$, of a perfect biomarker should be α while a

useless one is $0.5\alpha^2$. Similar to the probability explanation of AUC, the PAUC can be explained as the probability that the test result of a case is higher than that of a control with its value exceeding the corresponding $(1-\alpha)$th quantile value, i.e., $P(Y_i > Y_j > q | D_i = 1, D_j = 0), i \neq j$, where q satisfying $P(Y > q | D = 0) = \alpha$. Several estimation and inference procedures have been proposed by Emir, Wieand, Jung, and Ying (2000), Zhang, Zhou, Freeman, and Freeman (2002), and Dodd and Pepe (2003), among others. Since the AUC is a special case of the PAUC and all the inference procedures developed for the PAUC can be reduced to that for the AUC. A more thorough understanding of the ROC and the related measures can also be found in Zhou, McClish, and Obuchowski (2002) and Pepe (2003).

1.2 Motivating Examples

Two empirical examples are illustrated in this section to provide some perspectives that are not considered in the traditional ROC methodologies. As we will see in these studies that the disease status is induced from the time of a specific event, which is different from the traditional one. It motivates the development of novel methodologies in this thesis.

1.2.1 Angiography Coronary Artery Disease (CAD) Study

The first considered data were obtained from the British Columbia Vital Statistics database which has been analyzed by Lee, et al. (2006). A total of 1050 patients were recruited between 1993 and 1995 from two Vancouver teaching hospitals for selective coronary angiography. The blood of each recruited patient was taken and frozen to store. In 2002, the blood samples were thawed and the plasma biomarkers of CRP, SAA, IL-6, and tHcy were recorded. The disease time of patients were collected in 2004 and 95 patients were found to die by CAD during the study period. The aim of this study is mainly to seek the best plasma biomarker in classifying patient's CAD-related death status. The analysis results of Lee, et al. (2006) indicated that

the elevated levels of IL-6 and tHcy are superior than those of CRP and SAA. The authors concluded that the AUC of IL-6 is significantly higher than the others.

1.2.2 AIDS Clinical Trials Group (ACTG) 175 Study

In the ACTG 175 study, a total of 2467 HIV-1-infected patients, whose CD4 cell counts ranged from 200 to 500 cells cu/mm, were recruited between December 1991 and October 1992. Among these patients, 1395 patients received the prior antiretroviral therapy while the rest 1072 patients did not. During the study period, 308 patients died of all causes or were diagnosed with AIDS. In contrast to the CAD study, the biomarker CD4 cell counts were intermittently collected within the study period. These longitudinal measured biomarkers will provide updated information about the disease progress. Full details of the design, medical implication, and methods of this study can be found in Hammer et al. (1996).

1.3 Time-Dependent Monitoring Rule

Recent research in ROC methodology has extended the binary disease status D to the time-dependent setting where the time-dependent disease status is defined through the failure time of a specific event. This generalization further provide the time-evolution effect in classification. In the CAD study, the disease status is induced from the CAD-related death time in which a patient is treated as a case if he/she died by CAD before the time point of interest. In the same manner, the time to AIDS in the ACTG 175 study is used to define the disease status. To simplify the succeeding presentation, the failure time of a specific event is denoted by T. For any fixed time point t, the time-dependent disease status is classified as a case if $\{T \leq t\}$ and a control otherwise. Based on the time-dependent setting, Heagerty, Lumley, and Pepe (2000) first generalized the traditional $TPR(y)$ and $FPR(y)$ to $TPR_t(y) = P(Y > y|T \leq t)$ and $FPR_t(y) = P(Y > y|T > t)$, respectively. The time-dependent ROC, AUC, and PAUC are naturally derived and can be applied to

evaluate the performance of Y in classifying $\{T \leq t\}$ and $\{T > t\}$.

Instead of using a single biomarker, an appropriate combination of multiple ones will improve the accuracy in classifying subject's vital status. As shown in the CAD study the time-dependent AUCs of CRP, SAA, IL-6, and tHcy are not high enough. It is desirable to seek an optimal composite biomarker of them that can improve classification accuracy and possesses the highest time-dependent ROC curve at each time point. As to the ACTG 175 study, it has been well known that CD4 cell counts are highly correlated with AIDS. A further scientific question is to investigate the effects of other risk factors Z such as the baseline therapy status on the performance of the biomarker. For instance, does CD4 biomarker performs better for patients with prior antiretroviral therapy. The covariate-specific time-dependent AUC denoted by $\theta_t(Z_i, Z_j) = P(Y_i > Y_j | T_i \leq t, T_j > t, Z_i, Z_j)$ provides useful information to assess whether the prior therapy can enhance the classification ability of CD4 cell counts. Full understanding of this quantity could help searching suitable subpopulation where the biomarker performs well. A generalized linear regression model for $\theta_t(Z_i, Z_j)$ was usually considered in the research to account for the heterogeneity arising from Z.

1.4 Existing Methods for Time-Dependent Vital Status

By applying the Baye's rule, Heagerty, Lumley, and Pepe (2000) proposed nonparametric estimators for $TPR_t(y)$ and $FPR_t(y)$ under different censoring mechanisms. These authors also suggested to estimate the time-dependent AUC and PAUC via calculating the relevant trapezoidal area under the estimated time-dependent ROC curve. In contrast to the numerical integration technique, Chambless and Diao (2006) developed the recursive estimators for $TPR_t(y)$, $FPR_t(y)$, and the time-dependent AUC. Their recursive estimators can, however, only provide estimates on the observed failure times while practical research might be interested in the estimation of the time-dependent AUC at a pre-specified time point t. For direct estimation of

the time-dependent PAUC, there is no study in the literatures concerning this topic currently. To investigate the effects of the covariates on the performance of a test result, Cai et al. (2006) proposed a class of semiparametric regression models for the time-dependent true and false positive rates. The covariate-specific time-dependent AUC is then computed numerically while the effects of Z on $\theta_t(Z_i, Z_j)$ can not be well explained. In view of this point, a model for the relationship between covariates and the AUC becomes necessary. Under the time-invariant disease status, Dodd and Pepe (2003) proposed a semiparametric regression model for the covariate-specific AUC $P(Y_i > Y_j|D_i = 1, D_j = 0, Z_i, Z_j)$ and develop the corresponding estimation procedure for the parameters of interest. So far, there is still no statistical method for the covariate-specific time-dependent AUC.

As to the problem of seeking an optimal combination of multiple biomarkers $Y = (Y_1, \cdots, Y_p)$, existing methods mainly rely on appropriately modeling the relationship between T and Y. Provided $P(T \leq t|Y = y)$ is an increasing function of $g_t(y)$ for any fixed t, the transformation $g_t(Y)$ is derived to be the optimal composite biomarker in the sense that no other function of Y can have a higher time-dependent ROC curve. Research interest usually focuses on the estimation of such optimal transformation $g_t(\cdot)$. In applications, a Cox's proportional hazards model $\lambda(t|y) = \lambda_0(t)\exp(\beta^T y)$, where $\lambda_0(t)$ is the baseline hazard function and $\beta = (\beta_1, \cdots, \beta_p)^T$ is the effects of Y on the hazard rate, is widely used to find an optimal classifier. It is straightforwardly to see that $g_t(Y) = \beta^T Y$ in such model and β can be estimated by the partial likelihood estimation procedure. A more flexible model which also implies $g_t(Y) = \beta^T Y$ is the generalized accelerated failure time (GAFT) model $h(T) = -\beta^T Y + \varepsilon$. Here $h(\cdot)$ is an unknown and increasing link function and ε represents a random error with unknown distribution. To estimate β, Khan and Tamer (2007) and Song, Ma, Huang, and Zhou (2007) have proposed the partial rank (PR) estimator and the smoothed PR estimator, respectively. One merit of the PR estimation method is to handle marker-dependent censorship, which is more acceptable in practice. Without involving the censoring distribution in the estimation, Cai and Cheng (2008) provided

a more robust nonparametric estimator for β at the expense of assuming totally independent censorship. To incorporate the time influence nature of Y in classifying $\{T \leq t\}$ and $\{T > t\}$, Zheng, Cai, and Feng (2006) considered the logistic regression model $P(T \leq t|y) = \exp(\alpha_t + \beta_t^T y)/\{1 + \exp(\alpha_t + \beta_t^T y)\}$ with time-varying coefficient $\beta_t = (\beta_{t1}, \cdots, \beta_{tp})^T$. These authors showed that $g_t(Y) = \beta_t^T Y$ and applied the inverse probability weighting (IPW) technique to estimate β_t. Without further modeling of the censoring distribution, Chiang and Huang (2008) considered an imputation method to estimate the optimal composite biomarkers.

1.5 Main Contributions

With censored survival data, traditional methodologies for the ROC curve and the related summary indices cannot be applied directly. The first aim of this thesis is to propose some estimation procedures for the time-dependent AUC and PAUC under different censoring mechanisms. In Chapter 2 we find that the time-dependent PAUC can be expressed as a functional of $S(t, y) = P(T > t, Y > y)$, i.e., the joint survival function of T and Y. By substituting reasonable estimators for $S(t, y)$, a class of nonparametric estimators are derived. In contrast to the methods of Heagerty, Lumley, and Pepe (2000) and Chambless and Diao (2006), our estimators are easily computed without involving very complicated numerical calculation. As one can see that there is no explicit expressions for their estimators which might be difficult to develop the corresponding asymptotic properties. All of these authors suggested using the time-consuming bootstrap method to make statistical inferences. By applying the functional Delta method, our estimators are shown to converge weakly to Gaussian processes with estimated variance-covariance functions. The developed properties further facilitate us to construct the approximated pointwise and simultaneous confidence bands for the time-dependent AUC and PAUC.

In Chapter 3, we consider the generalized linear regression model

$$\theta_t(Z_i, Z_j) = h(\gamma_t^T Z_{ij}) \tag{1.1}$$

with time-varying coefficient $\gamma_t = (\gamma_{t1}, \cdots, \gamma_{tp})^T$ to model the relationship between the time-dependent AUC and covariates Z, where $h(\cdot)$ is a known smooth response function and Z_{ij} is a $p \times 1$ vector function of Z_i and Z_j. When data are completely observed, the method of Dodd and Pepe (2003) can be applied by simply replacing the binary disease status D by the time-dependent one $I(T \leq t)$. With the appearance of censoring, their method might lead to biased estimates without handling unavailable vital status. By generalizing the proposed nonparametric estimator for the time-dependent AUC, we develop an estimating equation for the parameter γ_t in this thesis. The large sample properties for the estimators of γ_t and $\theta_t(Z_i, Z_j)$ are also derived to form the basis of statistical inferences.

The issue of combining biomarkers to achieve higher classification ability is studied in Chapter 4. It can be seen that the usefulness of the existing methods mainly rely on the appropriate specification for the working model. Under the validity of well behaved model for failure time, the linear combination of biomarkers is optimal in classifying the vital status over time. Robust estimation procedure is required in practice with less restrictive model assumption. A more flexible extended generalized linear model (EGLM)

$$P(T \leq t|Y = y) = G_t(\beta_t^T y) \tag{1.2}$$

is used to characterize the relationship between T and Y, where $\beta_t = (\beta_{t1}, \cdots, \beta_{tp})^T$ is the p-variate time-varying parameter of interest and $G_t(\cdot)$ is an unknown strictly increasing function which may also depended on t. It is ensured from the increasing property of $G_t(\cdot)$ that $\beta_t^T Y$ is the optimal composite biomarker at time t. The flexibility of (1.2) can be seen that all the aforementioned models such as the GAFT model and the time-varying coefficient logistic regression model are special cases. Based on the EGLM, we propose estimating equations for β_t via maximizing the estimated time-dependent AUC quantities under different censoring mechanisms. Interestingly, the root-n consistency for estimators of β_t can be verified even if $G_t(\cdot)$ is treated as a nuisance parameter. Moreover, estimators for the time-dependent ROC curve and related summary measures of the optimal composite biomarker $\beta_t^T Y$ are also proposed

by applying the estimation methods in Chapter 2 with little modifications.

Chapter 2

Time-Dependent AUC and PAUC

As mentioned in Chapter 1, there is still no rigorous inference procedure for the time-dependent AUC and PAUC. Consider the censored survival data $\{X_i, \delta_i, Y_i\}_{i=1}^n$ with X_i being the minimum of failure time T_i and censoring time C_i, $\delta_i = I(X_i = T_i)$ denoting the censoring status, and Y_i being the biomarker of the ith subject. The aim of my research in this chapter is to propose easily computed nonparametric estimators for these classification accuracies and the related inference procedures. Here, we focus on the time-dependent PAUC $\theta_t(q_{\alpha t})$ with $FPR_t(y) \leq \alpha$, where $q_{\alpha t} = FPR_t^{-1}(\alpha) = \inf\{y : FPR_t(y) \leq \alpha\}$ is the $(1-\alpha)$th quantile of Y conditioning on $\{T > t\}$ at a fixed time point t. The reason for this is because the time-dependent PAUC with restricted $TPR_t(y)$ can be derived in the same manner by reversing the roles of cases and controls. In addition, the time-dependent AUC can be treated as a special case of $\theta_t(q_{\alpha t})$ by setting $\alpha = 1$. We show that the estimation procedures for $\theta_t(q_{\alpha t})$ are mainly based on nonparametric estimators of $S(t, y)$. The asymptotic Gaussian processes of the proposed estimators and the corresponding estimated variance-covariance functions facilitate the construction of inference procedures.

2.1 Nonparametric Estimation

From the geometric display of relevant $\{FPR_t(y), TPR_t(y)\}$ points in the marker space, $\theta_t(q_{\alpha t})$ has the form $-\int TPR_t(y) I(FPR_t(y) \leq \alpha) d_y FPR_t(y)$. Moreover, the conditional probabilities $FPR_t(y)$ and $TPR_t(y)$ can be expressed as a functional of $S(t, y)$:

$$FPR_t(y) = \frac{S(t, y)}{S(t, -\infty)} \text{ and } TPR_t(y) = \frac{S(0, y) - S(t, y)}{1 - S(t, -\infty)}. \tag{2.1}$$

It is straightforward to derive that

$$\theta_t(q_{\alpha t}) = \frac{-\int (S(0, u) - S(t, u)) I(u \geq q_{\alpha t}) d_u S(t, u)}{S(t, -\infty)(1 - S(t, -\infty))}. \tag{2.2}$$

In view of (2.2), the estimation problem of $\theta_t(q_{\alpha t})$ obviously becomes that of $S(t, y)$.

Under totally independent censoring (C is independent of (T, Y)), which is an appropriate assumption for the Type I censoring, two estimators

$$\widetilde{S}^{(C)}(t, y) = \frac{\sum_{i=1}^n I(X_i > t, Y_i > y)}{n \widetilde{S}_X(t)} \text{ and } \widetilde{S}^{(B)}(t, y) = \frac{\sum_{i=1}^n \delta_i I(X_i > t, Y_i > y)}{n \widetilde{S}_C(X_i)} \tag{2.3}$$

for $S(t, y)$ are proposed by Campbell (1981) and Burke (1988), respectively. The $\widetilde{S}_C(t)$ and $\widetilde{S}_X(t) = n^{-1} \sum_{i=1}^n I(X_i > t)$ are the Kaplan-Meier estimator of $S_C(t) = P(C > t)$ and an estimator of $S_X(t) = P(X > t)$. Using the estimators $\widetilde{S}^{(B)}(t, y)$ and $\widetilde{S}^{(C)}(t, y)$, we propose a more robust estimator against the violation of totally independent censoring as below.

$$\begin{aligned} \widetilde{\theta}_t(\widetilde{q}_{\alpha t}) &= \frac{-\int \{\widetilde{S}^{(B)}(0, u) - \widetilde{S}^{(B)}(t, u)\} I(u \geq \widetilde{q}_{\alpha t}) d_u \widetilde{S}^{(C)}(t, u)}{\{1 - \widetilde{S}^{(B)}(t, -\infty)\} \widetilde{S}^{(C)}(t, -\infty)} \\ &= \frac{1}{n^2} \sum_{i,j} \frac{\delta_i I(X_i \leq t, X_j > t) \phi_{ij}(\widetilde{q}_{\alpha t})}{\widetilde{S}_C(X_i) \widetilde{S}_X(t)(1 - \widetilde{S}_T(t))}, \end{aligned} \tag{2.4}$$

where $\phi_{ij}(y) = I(Y_i > Y_j \geq y)$ and $\widetilde{q}_{\alpha t} = \widetilde{FPR}_t^{-1}(\alpha)$ with $\widetilde{FPR}_t(y) = \widetilde{S}^{(C)}(t, y)/\widetilde{S}^{(C)}(t, -\infty)$. By substituting the time-dependent vital status $\{T \leq t\}$ for the binary disease status D, $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ will reduce to the the estimator of Dodd and Pepe (2003a) when the complete failure time data are available.

Since the censoring time might associate with the baseline biomarker, it can be verified that $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ is not a consistent estimator of $\theta_t(q_{\alpha t})$. Under a more flexible marker-dependent censoring assumption (Conditioning on Y, T and C are independent), Akritas (1994) suggested using $\widehat{S}(t,y) = -\int_y^\infty \widehat{S}_T(t|Y_i)d\widehat{S}_Y(y)$ to estimate $S(t,y)$, where

$$\widehat{S}_T(t|y) = \prod_{\{i:X_i\le t,\delta_i=1\}} \{1 - \frac{K_\lambda(\widehat{S}_Y(Y_i) - \widehat{S}_Y(y))}{n\widehat{S}_X(X_i|y)}\}, \tag{2.5}$$

$\widehat{S}_Y(y) = n^{-1}\sum_{j=1}^n I(Y_j > y)$, and $\widehat{S}_X(t|y) = n^{-1}\sum_{j=1}^n I(X_j \ge t)K_\lambda(\widehat{S}_Y(Y_j) - \widehat{S}_Y(y))$ are estimators for $S_T(t|y) = P(T > t|Y = y)$, $S_Y(y) = P(Y > y)$, and $S_X(t|y) = P(X > t|Y = y)$, respectively. As for the kernel function $K_\lambda(u) = \lambda^{-1}K(u/\lambda)$, a uniform density function $K(u) = 0.5I(|u| < 1)$ and a positive smoothing parameter λ are specified. Replacing $S(t,y)$ with $\widehat{S}(t,y)$ in (2.2), an alternative estimator for $\theta_t(q_{\alpha t})$ is proposed by

$$\begin{aligned}\widehat{\theta}_t(\widehat{q}_{\alpha t}) &= \frac{-\int(\widehat{S}(0,u) - \widehat{S}(t,u))I(u \ge \widehat{q}_{\alpha t})d_u\widehat{S}(t,u)}{\widehat{S}(t,-\infty)(1-\widehat{S}(t,-\infty))} \\ &= \frac{n^{-2}\sum_{i\ne j}(1-\widehat{S}_T(t|Y_i))\widehat{S}_T(t|Y_j)\phi_{ij}(\widehat{q}_{\alpha t})}{\widehat{S}_T(t)(1-\widehat{S}_T(t))},\end{aligned} \tag{2.6}$$

where $\widehat{S}_T(t) = \widehat{S}(t,-\infty)$ and $\widehat{q}_{\alpha t} = \widehat{FPR}_t^{-1}(\alpha)$ with $\widehat{FPR}_t(y) = \widehat{S}(t,y)/\widehat{S}_T(t)$. Note that the kernel function in (2.5) provides the nearest neighbor estimator (NNE) for $S_T(t|y)$. Other kernel functions are also possible and will yield different estimators. As mentioned in Akritas (1994), the asymptotic properties of $\widehat{S}(t,y)$ is irrelevant to the choice of kernel function under some regularity conditions and so is $\widehat{\theta}_t(\widehat{q}_{\alpha t})$. The author further showed that any other estimator for $S(t,y)$ is at least as dispersed as $\widehat{S}(t,y)$. One merit of the kernel function $K_\lambda(u)$ is that the choice of λ is irrelevant to the measurement scale of a biomarker.

When the research interest is focused on estimating the time-dependent PAUC with $\alpha < FPR_t(y) \le \alpha'$, $0 \le \alpha < \alpha' \le 1$, this classification accuracy can be expressed as $(\theta_t(q_{\alpha' t}) - \theta_t(q_{\alpha t}))$ and is estimated by $(\widetilde{\theta}_t(\widetilde{q}_{\alpha' t}) - \widetilde{\theta}_t(\widetilde{q}_{\alpha t}))$ or $(\widehat{\theta}_t(\widehat{q}_{\alpha' t}) - \widehat{\theta}_t(\widehat{q}_{\alpha t}))$

depending on the corresponding censoring structure. In the estimation of the time-dependent AUC, both $\phi_{ij}(\widetilde{q}_{\alpha t})$ and $\phi_{ij}(\widehat{q}_{\alpha t})$ will simplify to $I(Y_i > Y_j)$ and, hence, the estimators $\widetilde{\theta}_t(\widetilde{q}_{1t})$ and $\widehat{\theta}_t(\widehat{q}_{1t})$ are directly obtained. Let

$$\widetilde{TPR}_t(y) = \frac{\widetilde{S}^{(B)}(0,y) - \widetilde{S}^{(B)}(t,y)}{1 - \widetilde{S}^{(B)}(t,-\infty)} \text{ and } \widehat{TPR}_t(y) = \frac{\widehat{S}(0,y) - \widehat{S}(t,y)}{1 - \widehat{S}_T(t)}. \tag{2.7}$$

One can see that $\widetilde{\theta}_t(\widetilde{q}_{1t})$ and $\widehat{\theta}_t(\widehat{q}_{1t})$ can also be computed as the area under the right-continuous step function with jumps based on $(\widetilde{FPR}_t(y), \widetilde{TPR}_t(y))$ points and $(\widehat{FPR}_t(y), \widehat{TPR}_t(y))$ points, respectively. Interestingly, if we replace $I(Y_i > Y_j)$ by $I(Y_i \geq Y_j)$, the modified estimators are equivalent to compute the area under the left-continuous step function with jumps base on those estimates for $(FPR_t(y), TPR_t(y))$. In many applications, a most widely used approach to estimate the time-dependent AUC is to compute the area under the polygon formed by connecting the estimates $(FPR_t(y), TPR_t(y))$ points. We further find the derived estimator is merely to replace $I(Y_i > Y_j)$ by $I(Y_i > Y_j) + 0.5I(Y_i = Y_j)$. Since the differences among these estimators will converge to 0 with rate $O_p(1/n)$, we can conclude that the asymptotic properties of these modifications are the same. This fact provides a theoretical basis for the existing estimation methods of $\theta_t(q_{\alpha t})$ and facilitate the construction of inferences.

2.2 Asymptotic Properties

In this section, the weak convergence of the proposed nonparametric estimators are established. The asymptotic properties are derived via applying the functional central limit theorem to the independent and identically distributed (i.i.d.) representations of our estimators. These i.i.d. approximations can be further used to estimate the asymptotic variance-covariance functions and thus facilitate the development of statistical inference procedures for the time-dependent AUC and PAUC.

Theorem 2.1. Suppose that marker-dependent censoring and the conditions made

below and in Akritas (1994) are satisfied.

(A1) $f_t(y) = -\partial FPR_t(y)/\partial y$ exists with $\inf_t f_t(q_{\alpha t}) > 0$.

(A2) $\sup_t |\epsilon^{-1}\{FPR_t(q_{\alpha t}+\epsilon) - FPR_t(q_{\alpha t})\} + f_t(q_{\alpha t})| \to 0$ as $\epsilon \to 0$.

Then, $\sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t}))$ is uniformly approximated by $n^{-1/2}\sum_{i=1}^n \Psi_{\alpha i}(t)$, where $\Psi_{\alpha i}(t)$ is defined in the proof, and converges weakly to a Gaussian process in $t \in (0, \tau]$ with zero mean and variance-covariance function $\Sigma_\alpha(s,t) = E[\Psi_{\alpha i}(s)\Psi_{\alpha i}(t)]$, where τ satisfies $\inf_y P(X > \tau | Y = y) > 0$.

Proof. From Theorem 3.1 of Akritas (1994), one has

$$\sup_{t,y} |\sqrt{n}(\widehat{S}(t,y) - S(t,y)) - \frac{\sqrt{n}}{n}\sum_{i=1}^n V_i(t,y)| = o_p(1), \tag{2.8}$$

where $V_i(t,y) = \{S_T(t|Y_i) + \xi_i(t)\}I(Y_i > y) - S(t,y)$,

$$\xi_i(t) = -S_T(t|Y_i)\int_0^t \frac{d_u M_i(u|Y_i)}{S_X(u|Y_i)}, M_i(t|y) = I(X_i \le t)\delta_i + \ln S_T(t \wedge X_i|y),$$

and $t \wedge X_i = \min\{t, X_i\}$. Let $h_{ij}(t,y) = (1 - S_T(t|Y_i))S_T(t|Y_j)\phi_{ij}(y)$ with expectation $h(t,y)$, $\widehat{h}_{ij}(t,y) = (1 - \widehat{S}_T(t|Y_i))\widehat{S}_T(t|Y_j)\phi_{ij}(y)$, and $\widehat{h}(t,y) = n^{-2}\sum_{i\neq j}\widehat{h}_{ij}(t,y)$. The uniform consistency of $\widehat{S}_T(t|y)$ (cf. Dabrowska (1987)) ensures that

$$\begin{aligned}\widehat{h}(t,y) &= \frac{1}{n^2}\sum_{i\neq j} h_{ij}(t,y) + \frac{1}{n^2}\sum_{i\neq j}(S_T(t|Y_i) - \widehat{S}_T(t|Y_i))S_T(t|Y_j)\phi_{ij}(y) \\ &\quad + \frac{1}{n^2}\sum_{i\neq j}(1 - S_T(t|Y_i))(\widehat{S}_T(t|Y_j) - S_T(t|Y_j))\phi_{ij}(y) + r_{1n}(t,y) \end{aligned} \tag{2.9}$$

with $\sup_{t,y}|r_{1n}(t,y)| = o_p(n^{-1/2})$. By a direct calculation and (2.8), a simplified form of the second term in the righthand side of (2.9) is obtained as below.

$$\begin{aligned}&\frac{1}{n}\sum_{j=1}^n S_T(t|Y_j)I(Y_j > y)\{\frac{1}{n}\sum_{i=1}^n S_T(t|Y_i)I(Y_i > Y_j) - \widehat{S}(t,Y_j)\} \\ &\qquad = \frac{-1}{n^2}\sum_{i,j} S_T(t|Y_j)\xi_i(t)\phi_{ij}(y) + r_{2n}(t,y),\end{aligned} \tag{2.10}$$

where $\sup_{t,y}|r_{2n}(t,y)| = o_p(n^{-1/2})$. Similarly, the third term can be derived to be

$$\frac{1}{n^2}\sum_{i,j}(1 - S_T(t|Y_i))\xi_j(t)\phi_{ij}(y) + r_{3n}(t,y) \tag{2.11}$$

with $\sup_{t,y}|r_{3n}(t,y)| = o_p(n^{-1/2})$. It follows from (2.9)-(2.11), the decomposition of a U-statistic into a sum of degenerate U-statistics (Serfling (1980)), and Corollary 4 of Sherman (1994) that

$$\sup_{t,y}|\sqrt{n}(\widehat{h}(t,y) - h(t,y)) - \frac{\sqrt{n}}{n}\sum_{i=1}^{n} U_i(t,y)| = o_p(1) \tag{2.12}$$

with

$$U_i(t,y) = E[h_{ij}(t,y) + h_{ji}(t,y)|X_i, Y_i, \delta_i] - 2h(t,y) + \{S_Y(Y_i) - S(t,y)\}\xi_i(t)I(Y_i > y).$$

By the Taylor expansion of $\widehat{\theta}_t(y) = \widehat{h}(t,y)\{\widehat{S}_T(t)(1 - \widehat{S}_T(t))\}^{-1}$ at $(\widehat{h}(t,y), \widehat{S}_T(t)) = (h(t,y), S_T(t))$, (2.8), and (2.12), one has

$$\sup_{t,y}|\sqrt{n}(\widehat{\theta}_t(y) - \theta_t(y)) - \frac{\sqrt{n}}{n}\sum_{i=1}^{n}\frac{U_i(t,y) + \eta(t,y)V_i(t,-\infty)}{S_T(t)(1 - S_T(t))}| = o_p(1), \tag{2.13}$$

where $\eta(t,y) = h(t,y)(2S_T(t) - 1)\{S_T(t) - S_T^2(t)\}^{-1}$. Applying the functional central limit theorem, $\sqrt{n}(\widehat{\theta}_t(q_{\alpha t}) - \theta_t(q_{\alpha t}))$ is shown to converge to a mean zero Gaussian process in t.

As for the asymptotic Gaussian process of $\widehat{\theta}_t(\widehat{q}_{\alpha t})$, it can established through the equality $\sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t})) = \sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(\widehat{q}_{\alpha t})) + \sqrt{n}(\theta_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t}))$. Let $\sqrt{n}(\widehat{q}_{\alpha t} - q_{\alpha t}) = \sqrt{n}(Q(\widehat{S}) - Q(S))$ with $Q : S \to q_{\alpha t}$. By assumptions (A1)-(A2), the Hadamard differentiability of Q is a direct result of Lemma A.1 in Daouia, Florens, and Simar (2008). Together with the functional delta method (cf. Van der Vaart (2000)), we have

$$\sup_{t}|\sqrt{n}(\widehat{q}_{\alpha t} - q_{\alpha t}) - \frac{\sqrt{n}}{n}\sum_{i=1}^{n}\frac{V_i(t,q_{\alpha t}) - \alpha V_i(t,-\infty)}{f_t(q_{\alpha t})S_T(t)}| = o_p(1) \tag{2.14}$$

and the weak convergence of $\sqrt{n}(\widehat{q}_{\alpha t} - q_{\alpha t})$. Moreover, it is ensured by (a version of) Lemma 19.24 of van der Vaart (2000) and (2.13) that

$$\sup_{t}|\sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(\widehat{q}_{\alpha t}) - \sqrt{n}(\widehat{\theta}_t(q_{\alpha t}) - \theta_t(q_{\alpha t}))| = o_p(1). \tag{2.15}$$

By the first order Taylor expansion of $\theta_t(\widehat{q}_{\alpha t})$ at $\widehat{q}_{\alpha t} = q_{\alpha t}$, the continuity of $\partial\theta_t(y)/\partial y$, $\sup_t |\widehat{q}_{\alpha t} - q_{\alpha t}| = o_p(1)$, and the continuous mapping theorem, one derives that imply

$$\sup_t |\sqrt{n}(\theta_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t})) - \frac{\alpha S_T(t) - S_Y(q_{\alpha t})}{1 - S_T(t)} f_t(q_{\alpha t})\sqrt{n}(\widehat{q}_{\alpha t} - q_{\alpha t})| = o_p(1). \quad (2.16)$$

It follows from (2.13)-(2.16) that

$$\sup_t |\sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t})) - \frac{\sqrt{n}}{n}\sum_{i=1}^{n} \Psi_{\alpha i}(t)| = o_p(1), \text{ where} \quad (2.17)$$

$$\Psi_{\alpha i}(t) = \frac{U_i(t, q_{\alpha t}) + \eta(t, q_{\alpha t})V_i(t, -\infty) + (\alpha S_T(t) - S_Y(q_{\alpha t}))(V_i(t, q_{\alpha t}) - \alpha V_i(t, -\infty))}{S_T(t)(1 - S_T(t))}.$$

Finally, the proof is completed by applying the functional central limit theorem to the approximated term $n^{-1/2}\sum_{i=1}^{n} \Psi_{\alpha i}(t)$ in (2.17). □

The weak convergence of $\sqrt{n}(\widetilde{\theta}_t(\widetilde{q}_{\alpha t}) - \theta_t(q_{\alpha t}))$ can be derived in a similar manner and is given in the following theorem.

Theorem 2.2. Suppose that $S(t, y)$ and $S_C(t)$ are absolutely continuous. Under totally independent censoring, $\sqrt{n}(\widetilde{\theta}_t(\widetilde{q}_{\alpha t}) - \theta_t(q_{\alpha t}))$ is uniformly approximated by $n^{-1/2}\sum_{i=1}^{n} \Psi^*_{\alpha i}(t)$ and converges weakly to a Gaussian process in $t \in (0, \tau]$ with mean zero and variance-covariance function $\Sigma^*_\alpha(s, t) = E[\Psi^*_{\alpha i}(s)\Psi^*_{\alpha i}(t)]$, where $\Psi^*_{\alpha i}(t)$ is defined in the proof and τ satisfies $P(X > \tau) > 0$.

Proof. Let

$$g_{ij}(t, y) = \frac{\delta_i}{S_C(X_i)} I(X_i \le t, X_j > t)\phi_{ij}(y), g(t, y) = E[g_{ij}(t, y)],$$

$$\widetilde{g}_{ij}(t, y) = \frac{\delta_i}{\widetilde{S}_C(X_i)} I(X_i \le t, X_j > t)\phi_{ij}(y), \widetilde{g}(t, y) = \frac{1}{n^2}\sum_{i,j} \widetilde{g}_{ij}(t, y),$$

$$M_{Ti}(t) = \delta_i I(X_i \le t) + \ln S_T(t \wedge X_i), \text{ and } M_{Ci}(t) = (1 - \delta_i)I(X_i \le t) + \ln S_C(t \wedge X_i).$$

The martingale representation of survivor function (cf. Fleming and Harrington (1991)) ensures that $\sqrt{n}(\widetilde{S}_T(t) - S_T(t))$ and $\sqrt{n}(\widetilde{S}_C(t) - S_C(t))$ can be uniformly

approximated by

$$-S_T(t)\frac{\sqrt{n}}{n}\sum_{i=1}^{n}\int_0^t \frac{dM_{Ti}(u)}{S_X(u)} \text{ and } -S_C(t)\frac{\sqrt{n}}{n}\sum_{i=1}^{n}\int_0^t \frac{dM_{Ci}(u)}{S_X(u)}. \tag{2.18}$$

From (2.18), one can obtain that

$$\sup_{t\in[0,\tau]} |\sqrt{n}(\widetilde{g}(t,y)-g(t,y))-\{\frac{\sqrt{n}}{(n)_3}\sum_{i\neq j\neq k} g_{ij}(t,y)(1+\int_0^{X_i}\frac{dM_{Ck}(u)}{S_X(u)})-g(t,y)\}| = o_p(1). \tag{2.19}$$

By $\widetilde{\theta}_t(y) = \widetilde{g}(t,y)\{\widetilde{S}_X(t)(1-\widetilde{S}_T(t))\}^{-1}$ and taking the Taylor expansion of $\widetilde{\theta}_t(y)$ at $(\widetilde{g}(t,y),\widetilde{S}_X(t),\widetilde{S}_T(t)) = (g(t,y),S_X(t),S_T(t))$, it is further established that

$$\sup_{t,y} |\sqrt{n}(\widetilde{\theta}_t(y)-\theta_t(y)) - \frac{\sqrt{n}}{(n)_3}\sum_{i\neq j\neq k}\frac{U^*_{ijk}(t,y)}{S_X(t)(1-S_T(t))}| = o_p(1),$$

$$U^*_{ijk}(t,y) = g_{ij}(t,y)(1+\int_0^{X_i}\frac{dM_{Ck}(u)}{S_X(u)}) - g(t,y)(\frac{I(X_i>t)}{S_X(t)}+\frac{S_T(t)}{1-S_T(t)}\int_0^t\frac{dM_{Ti}(u)}{S_X(u)}). \tag{2.20}$$

Similar to the derivation of (2.12), $\sqrt{n}(\widetilde{\theta}_t(y)-\theta_t(y))$ is uniformly approximated by

$$\frac{\sqrt{n}}{n}\sum_{i=1}^{n}\frac{U^*_i(t,y)}{S_X(t)(1-S_T(t))}$$

with $U^*_i(t,y) = E[U^*_{ijk}(t,y)+U^*_{jik}(t,y)+U^*_{jki}(t,y)|X_i,Y_i,\delta_i]$. It follows from the functional central limit theorem that $\sqrt{n}(\widetilde{\theta}_t(q_{\alpha t})-\theta_t(q_{\alpha t}))$ converges weakly to a mean zero Gaussian process.

Similar to the arguments of (2.14) and (2.17), we have

$$\sup_{t,y}|\sqrt{n}(\widetilde{q}_{\alpha t}-q_{\alpha t}) - \frac{\sqrt{n}}{n}\sum_{i=1}^{n}\frac{(I(Y_i>q_{\alpha t})-\alpha)I(X_i>t)}{f_t(q_{\alpha t})S_X(t)}| = o_p(1). \tag{2.21}$$

This, the weak convergence of $\sqrt{n}(\widetilde{q}_{\alpha t}-q_{\alpha t})$ and

$$\sup_t|\sqrt{n}(\widetilde{\theta}_t(\widetilde{q}_{\alpha t})-\theta_t(q_{\alpha t})) - \frac{\sqrt{n}}{n}\sum_{i=1}^{n}\Psi^*_{\alpha i}(t)| = o_p(1) \text{ where} \tag{2.22}$$

with

$$\Psi^*_{\alpha i}(t) = \frac{U^*_i(t,q_{\alpha t})+(\alpha S_T(t)-S_Y(q_{\alpha t}))(I(Y_i>q_{\alpha t})-\alpha)I(X_i>t)}{S_X(t)(1-S_T(t))}.$$

can be ensured and, the proof is completed. □

Paralleling the proofs of Theorems 2.1-2.2, we can also derive the asymptotic Gaussian processes of $\sqrt{n}\{(\widehat{\theta}_t(\widehat{q}_{\alpha' t}) - \widehat{\theta}_t(\widehat{q}_{\alpha t})) - (\theta_t(q_{\alpha' t}) - \theta_t(q_{\alpha t}))\}$ and $\sqrt{n}\{(\widetilde{\theta}_t(\widetilde{q}_{\alpha' t}) - \widetilde{\theta}_t(\widetilde{q}_{\alpha t})) - (\theta_t(q_{\alpha' t}) - \theta_t(q_{\alpha t}))\}$ with the corresponding variance-covariance functions $E[(\Psi_{\alpha' i}(s) - \Psi_{\alpha i}(s))(\Psi_{\alpha' i}(t) - \Psi_{\alpha i}(t))]$ and $E[(\Psi^*_{\alpha' i}(s) - \Psi^*_{\alpha i}(s))(\Psi^*_{\alpha' i}(t) - \Psi^*_{\alpha i}(t))]$.

2.3 Inference Procedures

With the marker-dependent censorship, the variance-covariance function $\Sigma_\alpha(s,t)$ is suggested to be estimated by

$$\widehat{\Sigma}_\alpha(s,t) = \frac{1}{n}\sum_{i=1}^{n} \widehat{\Psi}_{\alpha i}(s)\widehat{\Psi}_{\alpha i}(t), \tag{2.23}$$

where

$$\widehat{\Psi}_{\alpha i}(t) = \frac{\widehat{U}_i(t,\widehat{q}_{\alpha t}) + \widehat{\eta}(t,\widehat{q}_{\alpha t})\widehat{V}_i(t,-\infty) + (\alpha\widehat{S}_T(t) - \widehat{S}_Y(\widehat{q}_{\alpha t}))(\widehat{V}_i(t,\widehat{q}_{\alpha t}) - \alpha\widehat{V}_i(t,-\infty))}{\widehat{S}_T(t)(1-\widehat{S}_T(t))},$$

$$\widehat{U}_i(t,y) = \frac{1}{n}\sum_{\{j:j\neq i\}} (\widehat{h}_{ij}(t,y) + \widehat{h}_{ji}(t,y)) - 2\widehat{h}(t,y) + (\widehat{S}_Y(Y_i) - \widehat{S}(t,y))\widehat{\xi}_i(t)I(Y_i > y),$$

$$\widehat{V}_i(t,y) = \{\widehat{S}_T(t|Y_i) + \widehat{\xi}_i(t)\}I(Y_i > y) - \widehat{S}(t,y), \widehat{\eta}(t,y) = \frac{\widehat{h}(t,y)(2\widehat{S}_T(t)-1)}{\widehat{S}_T(t) - \widehat{S}^2_T(t)},$$

$$\widehat{\xi}_i(t) = -\widehat{S}_T(t|Y_i)\int_0^t \frac{d_u\widehat{M}_i(u|Y_i)}{\widehat{S}_X(u|Y_i)}, \text{ and } \widehat{M}_i(t|y) = I(X_i \le t)\delta_i + \ln\widehat{S}_T(t\wedge X_i|y).$$

As to the setting of totally independent censoring, the variance-covariance function $\Sigma^*_\alpha(s,t)$ is estimated by

$$\widetilde{\Sigma}^*_\alpha(s,t) = \frac{1}{n}\sum_{i=1}^{n} \widetilde{\Psi}^*_{\alpha i}(s)\widetilde{\Psi}^*_{\alpha i}(t), \tag{2.24}$$

where

$$\widetilde{\Psi}^*_{\alpha i}(t) = \frac{\widetilde{U}^*_i(t,\widetilde{q}_{\alpha t}) + (\alpha\widetilde{S}_T(t) - \widehat{S}_Y(\widetilde{q}_{\alpha t}))(I(Y_i > \widetilde{q}_{\alpha t}) - \alpha)I(X_i > t)}{\widetilde{S}_X(t)(1-\widetilde{S}_T(t))},$$

$$\widetilde{U}^*_{ijk}(t,y) = \widetilde{g}_{ij}(t,y)(1 + \int_0^{X_i}\frac{d\widetilde{M}_{Ck}(u)}{\widetilde{S}_X(u)}) - \widetilde{g}(t,y)(\frac{I(X_i > t)}{\widetilde{S}_X(t)} + \frac{\widetilde{S}_T(t)}{1-\widetilde{S}_T(t)}\int_0^t \frac{d\widetilde{M}_{Ti}(u)}{\widetilde{S}_X(u)}),$$

$$\widetilde{U}_i^*(t,y) = \frac{1}{n^2} \sum_{\{j,k:j\neq i,k\neq i\}} (\widetilde{U}_{ijk}^*(t,y) + \widetilde{U}_{jik}^*(t,y) + \widetilde{U}_{jki}^*(t,y)),$$

$\widetilde{M}_{Ti}(t) = \delta_i I(X_i \leq t) + \ln \widetilde{S}_T(t \wedge X_i)$, and $\widetilde{M}_{Ci}(t) = (1-\delta_i) I(X_i \leq t) + \ln \widetilde{S}_C(t \wedge X_i)$.

Under the assumed conditions, the weak convergence of $\sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t}))$ and $\sqrt{n}(\widetilde{\theta}_t(\widetilde{q}_{\alpha t}) - \theta_t(q_{\alpha t}))$ to Gaussian processes are derived in Section 2.2. Together with the estimated variance-covariance matrices provided above, the inference procedure for $\theta_t(q_{\alpha t})$ can be established. A $(1-\varsigma)$, $0 < \varsigma < 1$, pointwise confidence interval for $\theta_t(q_{\alpha t})$ can be constructed via

$$\widehat{\theta}_t(\widehat{q}_{\alpha t}) \pm \frac{Z_{\varsigma/2}}{\sqrt{n}} \widehat{\Sigma}_\alpha^{1/2}(t,t) \text{ and } \widetilde{\theta}_t(\widetilde{q}_{\alpha t}) \pm \frac{Z_{\varsigma/2}}{\sqrt{n}} \widetilde{\Sigma}_\alpha^{*1/2}(t,t), \tag{2.25}$$

where $Z_{\varsigma/2}$ is the $(1-\varsigma/2)$ quantile value of the standard normal distribution. As for the simultaneous confidence band of $\theta_t(q_{\alpha t})$ within the subinterval $[\tau_1, \tau_2] \in [0, \tau]$ of interest, the re-sampling technique of Lin, Wei, Yang, and Ying (2000) and the i.i.d. representations in (2.17) and (2.22) can be used to determine critical points L_ς and L_ς^* so that

$$P(\sup_{t\in[\tau_1,\tau_2]} |\frac{\sqrt{n}(\widehat{\theta}_t(\widehat{q}_{\alpha t}) - \theta_t(q_{\alpha t}))}{\widehat{\Sigma}_\alpha^{1/2}(t,t)}| < L_\varsigma) \doteq 1-\varsigma \tag{2.26}$$

and

$$P(\sup_{t\in[\tau_1,\tau_2]} |\frac{\sqrt{n}(\widetilde{\theta}_t(\widetilde{q}_{\alpha t}) - \theta_t(q_{\alpha t}))}{\widetilde{\Sigma}_\alpha^{*1/2}(t,t)}| < L_\varsigma^*) \doteq 1-\varsigma. \tag{2.27}$$

Details of the re-sampling procedures are stated in the following steps:

1. Independently generate random samples $\{W_i^{(l)}\}_{i=1}^n$, $l = 1, \cdots, B$, from a standard normal distribution to calculate

$$\Gamma_l = \sup_{t\in[\tau_1,\tau_2]} |\frac{\sum_{i=1}^n W_i^{(l)} \widehat{\Psi}_{\alpha i}(t)}{\{n\widehat{\Sigma}_\alpha(t,t)\}^{1/2}}| \text{ and } \Gamma_l^* = \sup_{t\in[\tau_1,\tau_2]} |\frac{\sum_{i=1}^n W_i^{(l)} \widetilde{\Psi}_{\alpha i}(t)}{\{n\widetilde{\Sigma}_\alpha^*(t,t)\}^{1/2}}|.$$

2. Determine L_ς and L_ς^* to be the $(1-\varsigma)$ quantile of $\{\Gamma_l\}_{l=1}^B$ and $\{\Gamma_l^*\}_{l=1}^B$, respectively.

3. The approximated $(1-\varsigma)$ simultaneous confidence bands for $\{\theta_t(q_{\alpha t}) : t \in [\tau_1, \tau_2]\}$ are constructed via

$$\{\widehat{\theta}_t(\widehat{q}_{\alpha t}) \pm \frac{L_\varsigma}{\sqrt{n}}\widehat{\Sigma}_\alpha^{1/2}(t,t) : t \in [\tau_1, \tau_2]\} \text{ and } \{\widetilde{\theta}_t(\widetilde{q}_{\alpha t}) \pm \frac{L_\varsigma^*}{\sqrt{n}}\widetilde{\Sigma}_\alpha^{*1/2}(t,t) : t \in [\tau_1, \tau_2]\}. \tag{2.28}$$

Remark 1. From the proof in Section 2.2, both pointwise and simultaneous confidence bands for $(\theta_t(q_{\alpha' t}) - \theta_t(q_{\alpha t}))$ can also be constructed in a similar manner through the i.i.d. representations $n^{-1/2}\sum_{i=1}^n(\widehat{\Psi}_{\alpha' i}(t) - \widehat{\Psi}_{\alpha i}(t))$ or $n^{-1/2}\sum_{i=1}^n(\widetilde{\Psi}_{\alpha' i}(t) - \widetilde{\Psi}_{\alpha i}(t))$.

Chapter 3

Semiparametric Regression Model for the Time-Dependent AUC

Chapter 3 is devoted to explore the effects of covariates Z on the time-dependent classification accuracy of a biomarker. We consider a generalized linear regression model:

$$\theta_t(Z_i, Z_j) = h(\gamma_t^T Z_{ij}), i \neq j, \tag{3.1}$$

where $h(\cdot)$ is a known smooth link function, Z_{ij} is designed a $p \times 1$ vector function of Z_i and Z_j, and $\gamma_t = (\gamma_{t1}, \cdots, \gamma_{tp})^T$ is the vector time-varying coefficients of Z_{ij}. This model is used mainly to account for the heterogeneity arising from Z. The linear, logit, probit, complementary log, and complementary log-log models, for instance, are some potential choices for h. The logistic regression model has been shown to be effective in applications and is more natural when one wants to interpret the odds ratios for the categorical covariates. Motivated by the nonparametric estimation method for $\theta_t(q_{\alpha t})$, we propose estimating equations for γ_t based on the censored survival data $\{(X_i, \delta_i, Y_i, Z_i)\}_{i=1}^n$. Inference procedures for γ_t and $h(\gamma_t^T Z_{ij})$ are further provided in the succeeding section by applying the asymptotic Gaussian process property of estimator for γ_t.

3.1 Estimation

As we can see in (2.3) with $\alpha = 1$, $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ can be obtained by solving the following estimating equation

$$\sum_{i \neq j} \frac{\delta_i I(X_i \leq t, X_j > t)}{S_C(X_i) S_C(t)} (I(Y_i > Y_j) - \theta_t(q_{1t})) = 0, \tag{3.2}$$

with $S_C(t)$ being substituted by the Kaplan-Meier estimator $\widetilde{S}_C(t)$. When $S_C(t|z^*) = P(C > t|Z^* = z^*)$, $Z^* = (Y, Z)$, is known, we generalize the equation (3.2) to our data setting. By using the property

$$E[\frac{\delta_i I(X_i \leq t, X_j > t)}{S_C(X_i|Z_i^*) S_C(t|Z_j^*)} (I(Y_i > Y_j) - h(\gamma_t^T Z_{ij}))|Z_i, Z_j] = 0, i \neq j, \tag{3.3}$$

the estimators of γ_t is proposed to be the solution of the estimating equations

$$\begin{aligned} U_t(\gamma, S_C) &= \frac{1}{(n)_2} \sum_{i \neq j} Z_{ij} \frac{h'(\gamma^T Z_{ij})}{\nu(\gamma^T Z_{ij})} \frac{\delta_i I(X_i \leq t, X_j > t)}{S_C(X_i|Z_i^*) S_C(t|Z_j^*)} (I(Y_i > Y_j) - h(\gamma^T Z_{ij})) \\ &\triangleq \frac{1}{(n)_2} \sum_{i \neq j} U_{tij}(\gamma, S_C), \end{aligned} \tag{3.4}$$

where h' is the derivative of h, $\nu = h(1-h)$ represents the conditional variance of $I(Y_i > Y_j)$, and $(n)_m = n(n-1)\cdots(n-m+1)$. We note that (3.3) holds under the validity of (3.1) and conditional independent censoring (Conditioning on Z^*, T and C are independent). Generally, the range of h should be restricted within the interval $(0, 1)$ to ascertains that ν is bounded away from zero.

Since $S_C(t|z^*) = P(C > t|Z^* = z^*)$ is usually unknown, an appropriate consistent estimator should provided. To avoid the complicated and mathematical intractable modeling between C and Z^*, a widely used Cox's proportional hazards model can be applied. A practical estimation approach for $S_C(t|z^*)$ is based on the partial likelihood estimation procedure and the Breslow estimator (Fleming and Harrington (1991)). Another option is to use an accelerated failure time (AFT) model in the current study. When C is further independent of Z^*, $S_C(t|z^*)$ can be reduced to $S_C(t)$ and the Kaplan-Meier estimator is suggested. By substituting a consistent

estimator $\widehat{S}_C(t|z^*)$ for $S_C(t|z^*)$, we propose to estimate γ_t by the solution $\widehat{\gamma}_t$ of the equation $U_t(\widehat{\gamma}_t, \widehat{S}_C) = 0$. The covariate-specific time-dependent AUC at $Z_{ij} = z_0$ is naturally estimated by $h(\widehat{\gamma}_t^T z_0)$.

3.2 Asymptotic Properties

In this section, the large sample properties of $\widehat{\gamma}_t$ are considered in the interval $[\varsigma_0, \varsigma_1]$ with $\varsigma_0 = \inf_u\{u : \sup_{z^*} P(T > u|Z^* = z^*) < 1\}$ and $\varsigma_1 = \sup_u\{u : \inf_{z^*} P(X > u|Z^* = z^*) > 0\}$. For the convenience of succeeding presentation, let $\|\cdot\|$ denote the supremum norm of a vector or a square matrix and

$$H(t) = -E[Z_{ij}Z_{ij}^T \frac{(h^{'}(\gamma_t^T Z_{ij}))^2 \delta_i I(X_i \leq t, X_j > t)}{\nu(\gamma_t^T Z_{ij}) S_C(X_i|Z_i^*) S_C(t|Z_j^*)}].$$

Some mild conditions are further assumed in the following:

(A1) $h(u)$ is monotone, and $h(u)$ and $h^{'}(u)$ are Lipschitz continuous and bounded.

(A2) Z is bounded.

(A3) $H(t)$ is nonsingular for $t \in [\varsigma_0, \varsigma_1]$.

(A4) γ_t is cadlag on $[\varsigma_0, \varsigma_1]$.

(A5) $\sup_{t,z^*} |\widehat{S}_C(t|z^*) - S_C(t|z^*) - n^{-1}\sum_{i=1}^n G_i(t, z^*)| = o_p(n^{-1/2})$ with $G_i(t, z^*)$ being a zero mean function of $(X_i, \delta_i, Y_i, Z_i)$ for any fixed (t, z^*).

Although the bounded assumption in (A2) is frequently occurring in many empirical examples, this condition can be relaxed via making assumption on the moments of Z. It is entailed from assumptions (A1)-(A4) that the classes of kernel functions of $U_t(\gamma, S_C)$ and $\partial U_t(\gamma, S_C)/\partial\gamma$ indexed by (γ, t) are Euclidean. This is a necessary condition in the application of uniform consistency of U-process and the functional central limit theorem. Note that assumption (A5) is automatically satisfied for the survivor function estimated from the proportional hazards model and the Kaplan-Meier estimator under suitable conditions. For the Kaplan-Meier estimator, the martingale

representation in (2.18) shows that $G_i(t, z^*) = -S_C(t) \int_0^t dM_{Ci}(u)/S_X(u)$.

Theorem 3.1. Suppose that assumptions (A1)-(A5) are satisfied. Under the validity that T and C are independent conditioning on Z^*, $\sup_{t\in[\varsigma_0,\varsigma_1]} \parallel \widehat{\gamma}_t - \gamma_t \parallel = o_p(1)$ and $\sqrt{n}(\widehat{\gamma}_t - \gamma_t)$ is uniformly approximated by $n^{-1/2}\sum_{i=1}^n -\{H(t)\}^{-1}U_i(t)$ and converges weekly to a Gaussian process in $t \in [\varsigma_0, \varsigma_1]$ with mean zero and variance-covariance function $\Sigma(s,t) = H^{-1}(s)E[U_i(s)U_i^T(t)]H^{-1}(t)$, where $U_i(t)$ is defined in the proof.

Proof. We first show the uniform consistency of $\widehat{\gamma}_t$. Using the property $E[U_t(\gamma_t, S_C)] = 0$ and Corollary 7 of Sherman (1994), one has

$$\sup_{t\in[\varsigma_0,\varsigma_1]} \parallel \frac{\partial}{\partial\gamma} U_t(\gamma_t, S_C) - H(t) \parallel = o_p(1). \tag{3.5}$$

It follows from (3.5) and assumption (A3) that γ_t is the unique solution of $E[U_t(\gamma, S_C)] = 0$. By assumption (A5), the uniform consistency of $\widehat{\gamma}_t$ can be derived by verifying

$$\sup_{\gamma,t} \parallel U_t(\gamma, S_C) - E[U_t(\gamma, S_C)] \parallel = o_p(1). \tag{3.6}$$

Since the property in (3.6) is a direct consequence of the uniform convergence theorem of a U-process, it is obviously to obtain that $\sup_{t\in[\varsigma_0,\varsigma_1]} \parallel \widehat{\gamma}_t - \gamma_t \parallel = o_p(1)$.

As to the weak convergence, the first order Taylor expansion of $U_t(\widehat{\gamma}_t, \widehat{S}_C)$ with respect to γ_t implies that

$$0 = \sqrt{n}U_t(\widehat{\gamma}_t, \widehat{S}_C) = \sqrt{n}U_t(\gamma_t, \widehat{S}_C) + \{\frac{\partial}{\partial\gamma_t} U_t(\widehat{\gamma}_t^*, \widehat{S}_C)\}\sqrt{n}(\widehat{\gamma}_t - \gamma_t), \tag{3.7}$$

where $\widehat{\gamma}_t^*$ lies on the line segment between $\widehat{\gamma}_t$ and γ_t. From (3.5), assumption (A5), and the uniform consistency of $\widehat{\gamma}_t$ to γ_t, one can obtain that

$$\sup_{t\in[\varsigma_0,\varsigma_1]} \parallel \frac{\partial}{\partial\gamma} U_t(\widehat{\gamma}_t^*, \widehat{S}_C) - H(t) \parallel = o_p(1). \tag{3.8}$$

It is further implied from assumption (A5) that

$$\sup_{t\in[\varsigma_0,\varsigma_1]} \parallel \sqrt{n}U_t(\gamma_t, \widehat{S}_C) - \frac{\sqrt{n}}{(n)_4} \sum_{i\neq j\neq k\neq l} U_{ijkl}(t) \parallel = o_p(1), \tag{3.9}$$

where

$$U_{ijkl}(t) \;=\; U_{tij}(\gamma_t, S_C)\{1-\frac{G_k(X_i, Z_i^*)}{S_C(X_i|Z_i^*)}\}\{1-\frac{G_l(t, Z_j^*)}{S_C(t|Z_j^*)}\}.$$

Together with (3.7)-(3.8), the decomposition of the above U-statistic into the sum of degenerate U-statistics, and Corollary 4 of Sherman (1994), the weak convergence of $\sqrt{n}(\widehat{\gamma}_t - \gamma_t)$ is ensured through the following random quantity:

$$\frac{\sqrt{n}}{n}\sum_{i=1}^{n} -\{H(t)\}^{-1}U_i(t), \tag{3.10}$$

where $U_i(t) = E[U_{ijkl}(t) + U_{jikl}(t) + U_{jkil}(t) + U_{jkli}(t)|X_i, Y_i, Z_i, \delta_i]$ for $i \neq j \neq k \neq l$. By a direct calculation and assumption (A5), $U_i(t)$ can be further derived as

$$U_i(t) = E[U_{tij}(\gamma_t, S_C)+U_{tji}(\gamma_t, S_C)-U_{tjk}(\gamma_t, S_C)\{\frac{G_i(X_j, Z_j^*)}{S_C(X_j|Z_j^*)}+\frac{G_i(t, Z_k^*)}{S_C(t|Z_k^*)}\}|X_i, Y_i, Z_i, \delta_i].$$

Applying the functional central limit theorem, Theorem 3.1 is obtained. □

The weak convergence of the covariate-specific time-dependent AUC estimator can be ensured from Theorem 3.1. From (3.10) and using the functional Delta method, $\sqrt{n}(h(\widehat{\gamma}_t^T z_0) - h(\gamma_t^T z_0))$ is uniformly approximated by

$$\frac{\sqrt{n}}{n}\sum_{i=1}^{n} -h'(\gamma_t^T z_0)z_0^T\{H(t)\}^{-1}U_i(t). \tag{3.11}$$

Similarly, $\sqrt{n}(h(\widehat{\gamma}_t^T z_0) - h(\gamma_t^T z_0))$ can be shown to converge weakly to a Gaussian process in t with mean zero and variance-covariance matrix

$$\Gamma(s,t) = h'(\gamma_s^T z_0)h'(\gamma_t^T z_0)z_0^T\Sigma(s,t)z_0.$$

3.3 Inference Procedures

For the asymptotic variance-covariance matrix $\Sigma(s,t)$, a sandwich-type estimator is proposed. First, the quantity $E[U_i(s)U_i^T(t)]$ is estimated by the sample analogue $n^{-1}\sum_{i=1}^{n}\widehat{U}_i(s)\widehat{U}_i^T(t)$ with

$$\widehat{U}_i(t) = \frac{1}{n^2}\sum_{j,k} U_{tij}(\widehat{\gamma}_t, \widehat{S}_C) + U_{tji}(\widehat{\gamma}_t, \widehat{S}_C) - U_{tjk}(\widehat{\gamma}_t, \widehat{S}_C)\{\frac{\widehat{G}_i(X_j, Z_j^*)}{\widehat{S}_C(X_j|Z_j^*)} + \frac{\widehat{G}_i(t, Z_k^*)}{\widehat{S}_C(t|Z_k^*)}\}$$

and $\widehat{G}_i(t, z^*)$ being a consistent estimator of $G_i(t, z^*)$. By using $\widehat{H}(t) = \partial U_t(\widehat{\gamma}_t, \widehat{S}_C)/\partial\gamma$ as an estimator of $H(t)$, the estimator of $\Sigma(s,t)$ is given by

$$\widehat{\Sigma}(s,t) = \{\widehat{H}(s)\}^{-1}\{\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_i(s)\widehat{U}_i^T(t)\}\{\widehat{H}(t)\}^{-1}. \tag{3.12}$$

Same with the foregoing argument, the Kaplan-Meier estimator is applied when $S_C(t|z^*) = S_C(t)$ and, hence, $\widehat{G}_i(t, z^*) = -\widetilde{S}_C(t)\int_0^t d\widetilde{M}_{Ci}(u)/\widetilde{S}_X(u)$.

By applying the asymptotic properties established in Section 3.2, the approximated $(1-\varsigma)$ confidence region of γ_t are given by

$$\{\gamma_t : n(\widehat{\gamma}_t - \gamma_t)^T\{\widehat{\Sigma}(t,t)\}^{-1}(\widehat{\gamma}_t - \gamma_t) \le \chi^2_{p,1-\varsigma}\}, \tag{3.13}$$

where $\chi^2_{p,1-\varsigma}$ is the $100(1-\varsigma)$th percentile of the Chi-square distribution with p degrees of freedom. The simultaneous confidence band for γ_t over a pre-specified time period $\Upsilon \subset [\varsigma_0, \varsigma_1]$ can also be constructed by the re-sampling technique. Details are stated as below.

1. Independently generate random samples $\{W_i^{(b)}\}_{i=1}^n$, $b = 1, \cdots, B$, from a standard normal distribution to calculate

$$\Gamma_b = \sup_{t\in\Upsilon}\sum_{i=1}^{n} W_i^{(b)2}\widehat{U}_i^T(t)\{\widehat{H}(t)\}^{-1}\{n\widehat{\Sigma}(t,t)\}^{-1/2}\{\widehat{H}(t)\}^{-1}\widehat{U}_i(t).$$

2. Determine L_ς to be the $(1-\varsigma)$ quantile of $\{\Gamma_b\}_{b=1}^B$.

3. The approximated $(1-\varsigma)$ simultaneous confidence band for $\{\gamma_t : t \in \Upsilon\}$ is constructed via

$$\{\{\gamma_t : n(\widehat{\gamma}_t - \gamma_t)^T\{\widehat{\Sigma}(t,t)\}^{-1}(\widehat{\gamma}_t - \gamma_t) \le L_\varsigma\} : t \in \Upsilon\}. \tag{3.14}$$

As to the covariate-specific time-dependent AUC $h(\gamma_t^T z_0)$ for any fixed z_0, an approximated $(1-\varsigma)$ confidence interval is constructed to be

$$h(\widehat{\gamma}_t^T z_0) \pm \frac{z_{\varsigma/2}}{\sqrt{n}}h'(\widehat{\gamma}_t^T z_0)\{z_0^T\widehat{\Sigma}(t,t)z_0\}^{1/2}. \tag{3.15}$$

Similarly, the simultaneous confidence band can be constructed through the above Step 1-3 except the quantity we calculate in Step 1 is replace by

$$\sup_{t\in\Upsilon}|\sum_{i=1}^{n}\frac{W_i^{(b)}z_0^T\{\widehat{H}(t)\}^{-1}\widehat{U}_i(t)}{\{nz_0\widehat{\Sigma}(t,t)z_0^T\}^{1/2}}|,$$

and the simultaneous confidence band for $\{h(\gamma_t^T z_0) : t \in \Upsilon\}$ is established as

$$\{h(\widehat{\gamma}_t^T z_0) \pm \frac{L_\varsigma}{\sqrt{n}}h'(\widehat{\gamma}_t^T z_0)\{z_0^T\widehat{\Sigma}(t,t)z_0\}^{1/2} : t \in \Upsilon\}. \tag{3.16}$$

Chapter 4

Optimal Composite Markers

As indicated in the analysis results of Chiang and Huang (2009) for the CAD study, the time-dependent AUC of each marker, CRP, SAA, IL-6, and tHcy, is generally not large enough to classify the CAD-related vital status over time. In order to enhance the classification power, we aim to seek the optimal composite biomarkers based on a flexible extended generalized linear model (EGLM)

$$P(T \leq t|Y = y) = G_t(\beta_t^T y), \tag{4.1}$$

where $Y = (Y_1, \cdots, Y_p)^T$ is a vector of continuous biomarkers measured at or before the outset of study, $\beta_t = (\beta_{t1}, \cdots, \beta_{tp})^T$ is the vector of time-dependent coefficients, and $G_t(\cdot)$ is an unknown link function and increasing in the linear predictor for each time point t. Let $ROC_t(\beta_t)$ be the time-dependent ROC curve of $\beta_t^T Y$ which displays the pair values of $FPR_t(c, \beta_t) = P(\beta_t^T Y > c|T > t)$ and $TPR_t(c, \beta_t) = P(\beta_t^T Y > c|T \leq t)$ for varying threshold value c. Using Neyman-Pearson lemma, we can show that $ROC_t(\beta_t)$ is higher than any ROC curve of the transformation of Y under the validity of model (4.1). It means that the linear predictor $\beta_t^T Y$ is the optimal composite biomarker in classifying patients who survive at t or not and thus the corresponding time-dependent AUC $\theta_t(\beta_t)$ is thus the greatest.

Under independent and marker-dependent censoring mechanisms, two types of

objective functions ate proposed for the estimation of β_t based on the censored survival data $\{(X_i, \delta_i, Y_i)\}_{i=1}^n$. Moreover, the estimators for $RTOC_t(\beta_t)$ and classification accuracy $\theta_t(\beta_t)$ are provided. It is found that β_t is not identifiable for the unspecified link function but $RTOC_t(\beta_t) = RTOC_t(a\beta_t)$ any positive constant a. The issue of non-identifiability can be circumvented by imposing a reasonable constraint on the parameters. Without loss of generality, β_{t1} is set to be one provided that the marker Y_1 is associated with T. Thus, the optimality will be reduced to search for the true parameters $(\beta_{t2}, \cdots, \beta_{tp})^T$ in the $(p-1)$-dimensional compact parameter space $\mathcal{B}_t$. For the convenience of succeeding presentation, we let $\beta_t = (\beta_{t2}, \cdots, \beta_{tp})^T$ and define $\beta_t^T Y = Y_1 + \beta_{t2} Y_2 + \cdots + \beta_{tp} Y_p$.

4.1 Estimation Procedures

Sine

$$\beta_t = \arg\max_{\beta} \theta_t(\beta), \tag{4.2}$$

the target function is designed as a sample analogue of $\theta_t(\beta)$ for the estimation of β_t. Under totally independent censoring, an estimator of $\theta_t(\beta)$ for given β is obtained, by substituting $I(\beta^T Y_i > \beta^T Y_j)$ for $I(Y_i > Y_j)$ in (2.4) with $\alpha = 1$, as

$$\widetilde{\theta}_t(\beta) = \frac{1}{(n)_2} \sum_{i \neq j} \frac{\delta_i I(X_i \leq t, X_j > t, \beta^T Y_i > \beta^T Y_j)}{\widetilde{S}_C(X_i)\widetilde{S}_X(t)(1 - \widetilde{S}_T(t))}. \tag{4.3}$$

Thus, β_t is suggested to be estimated by $\widetilde{\beta}_t$, which is the maximizer of $\widetilde{\theta}_t(\beta)$. One may further assess the performance of $\beta_t^T Y$ through $RTOC_t(\beta_t)$ and the classification accuracy $\theta_t(\beta_t)$. By generalizing the estimators $\widetilde{TPR}_t(y)$ and $\widetilde{FPR}_t(y)$ in Section 2.1 to our data setting, $TPR_t(c, \beta_t)$, $FPR_t(c, \beta_t)$, and $\theta_t(\beta_t)$ are naturally estimated by

$$\widetilde{TPR}_t(c, \widetilde{\beta}_t) = \sum_{i=1}^n \frac{\delta_i I(X_i \leq t, \widetilde{\beta}_t^T Y_i > c)}{n\widetilde{S}_C(X_i)(1 - \widetilde{S}_T(t))}, \widetilde{FPR}_t(c, \widetilde{\beta}_t) = \sum_{i=1}^n \frac{I(X_i > t, \widetilde{\beta}_t^T Y_i > c)}{n\widetilde{S}_X(t)}, \tag{4.4}$$

and $\widetilde{\theta}_t(\widetilde{\beta}_t)$, respectively.

In the CAD study, the censoring times of patients might be affected by the baseline plasma biomarkers. The assumption of marker-dependent censoring should be more reasonable and $\widetilde{\beta}_t$ might be biased. Let $S_X(t,y) = P(X > t|Y = y)f_Y(y)$ and $S_{\delta X}(t,y) = P(X > t, \delta = 1|Y = y)f_Y(y)$ with $f_Y(y)$ being the probability density function of Y. These bivariate functions are estimated by the smoothing estimators

$$\widehat{S}_X(t,y) = \sum_{i=1}^{n} \frac{I(X_i \geq t)K_\lambda(Y_i - y)}{n} \text{ and } \widehat{S}_{\delta X}(t,y) = \sum_{i=1}^{n} \frac{\delta_i I(X_i \geq t)K_\lambda(Y_i - y)}{n},$$

where $K_\lambda(u) = \prod_{j=1}^{p} \lambda_j^{-1} K(u_j \lambda_j^{-1})$, $u = (u_1, \cdots, u_p)^T$, $\lambda = (\lambda_1, \cdots, \lambda_p)^T$ is a vector of non-negative smoothing parameters, and $K(v)$ is symmetric about zero with bounded variation and satisfies $\int K(v)dv = 1$, $\int v^k K(v) = 0$, $k = 1, \cdots, p$, and $\int v^{p+1}K(v) < \infty$. The reason of using a higher order kernel function $K(u)$ is mainly to ensure $\sqrt{n}$-consistency of the proposed estimator. From (2.6) with $\alpha = 1$, an alternative objective function

$$\widehat{\theta}_t(\beta) = \frac{1}{(n)_2} \sum_{i \neq j} \frac{(1 - \widehat{S}_T(t|Y_i))\widehat{S}_T(t|Y_j)I(\beta^T Y_i > \beta^T Y_j)}{\widehat{S}_T(t)(1 - \widehat{S}_T(t))} \tag{4.5}$$

is considered, where

$$\widehat{S}_T(t|y) = \mathcal{P}_0^t\{1 + \frac{d_u \widehat{S}_{\delta X}(u,y)}{\widehat{S}_X(u,y)}\} = \prod_{\{i:X_i \leq t, \delta_i = 1\}} \{1 - \frac{K_\lambda(Y_i - y)}{n\widehat{S}_X(X_i, y)}\}$$

is a smoothing estimator of $S_T(t|y)$, $\widehat{S}_T(t) = n^{-1}\sum_{i=1}^{n} \widehat{S}_T(t|Y_i)$ is an estimator of $S_T(t)$, and $\mathcal{P}_0^t$ denotes the infinite product integral over $[0,t]$. Together with (4.2), an estimator $\widehat{\beta}_t$ of β_t is defined as the maximizer of $\widehat{\theta}_t(\beta)$. As for the quantities quantities $TPR_t(c,\beta_t)$, $FPR_t(c,\beta_t)$ and $\theta_t(\beta_t)$, we propose the corresponding estimators

$$\widehat{TPR_t}(c,\widehat{\beta}_t) = \sum_{i=1}^{n} \frac{(1 - \widehat{S}_T(t|Y_i))I(\widehat{\beta}_t^T Y_i > c)}{n\{1 - \widehat{S}_T(t)\}}, \widehat{FPR_t}(c,\widehat{\beta}_t) = \sum_{i=1}^{n} \frac{\widehat{S}_T(t|Y_i)I(\widehat{\beta}_t^T Y_i > c)}{n\widehat{S}_T(t)}, \tag{4.6}$$

and $\widehat{\theta}_t(\widehat{\beta}_t)$.

We note that the optimality of $\beta_t^T Y$ is mainly based on the increasing property of link function. The facts of (4.2) motivates the estimation of β_t via maximizing

$\widetilde{\theta}_t(\beta)$ and $\widehat{\theta}_t(\beta)$. Provided that (4.2) is valid, the linear predictor $\beta_t^T Y$ is still optimal in the sense that $\theta_t(\beta_t)$ is the largest among the class of linear combinations of Y. The asymptotic properties of $\widetilde{\beta}_t$ and $\widehat{\beta}_t$ derived in Section 4.2 are applicable when β_t is the unique maximizer of $\theta_t(\beta)$. The monotonicity of $\widetilde{TPR}_t(c, \widetilde{\beta}_t)$, $\widetilde{FPR}_t(c, \widetilde{\beta}_t)$, $\widehat{TPR}_t(c, \widehat{\beta}_t)$, and $\widehat{FPR}_t(c, \widehat{\beta}_t)$ in c further imply the monotonicity of the estimators for $ROC_t(\beta_t)$. Similar to the discussion in Section 2.1, the values of $\widetilde{\theta}_t(\widetilde{\beta}_t)$ and $\widehat{\theta}_t(\widehat{\beta}_t)$ are exactly the areas under the right-continuous step function with jumps based on the estimated $(\widetilde{FPR}_t(c, \widetilde{\beta}_t), \widetilde{TPR}_t(c, \widetilde{\beta}_t))$ points and $(\widehat{FPR}_t(c, \widehat{\beta}_t), \widehat{TPR}_t(c, \widehat{\beta}_t))$ points, respectively. Different modifications of $I(\beta^T Y_i > \beta^T Y_j)$ will result in different geometric meanings but the same large sample properties.

The implementation of the optimization procedures are time consuming due to the non-differentiability an indicator function. To overcome this problem, the distribution-like kernel functions of Horowitz (1992) can be applied to approximate $I(\beta^T Y_i > \beta^T Y_j)$. By incorporating our theoretical results, the same large sample properties of the smoothed nonparametric estimators can be derived similarly as in Song et al. (2007). In our numerical studies, $\widehat{S}_T^*(t|y) = \widehat{S}_T(t|y) I(0 \le \widehat{S}_T(t|y) \le 1) + I(\widehat{S}_T(t|y) > 1)$ is used to substitute for $\widehat{S}_T(t|y)$ in estimation and modifies unreasonable conditional survival estimates caused by a higher order kernel smoother.

4.2 Asymptotic Properties

It can be shown that the target function are

$$\widetilde{\theta}_t(\beta) = \frac{\widetilde{H}_t^{(1)}(\beta)}{\widetilde{S}_X(t)(1 - \widetilde{S}_T(t))} \text{ and } \widehat{\theta}_t(\beta) = \frac{\widehat{H}_t^{(2)}(\beta) - (n)_2^{-1} \sum_{i<j} \widehat{S}_T(t|Y_i)\widehat{S}_T(t|Y_j)}{\widehat{S}_T(t)(1 - \widehat{S}_T(t))} \quad (4.7)$$

with

$$\widetilde{H}_t^{(1)}(\beta) = \frac{1}{(n)_2} \sum_{i \neq j} \frac{\delta_i}{\widetilde{S}_C(X_i)} I(X_i \le t, X_j > t, \beta^T Y_i > \beta^T Y_j) \quad (4.8)$$

and

$$\widehat{H}_t^{(2)}(\beta) = \frac{1}{(n)_2} \sum_{i \neq j} \widehat{S}_T(t|Y_j) I(\beta^T Y_i > \beta^T Y_j). \quad (4.9)$$

Thus, the asymptotic properties of $\widetilde{\beta}_t$ and $\widehat{\beta}_t$ can be established based on the facts:

$$\widetilde{\beta}_t = \arg\max_{\beta} \widetilde{H}_t^{(1)}(\beta) \text{ and } \widehat{\beta}_t = \arg\max_{\beta} \widehat{H}_t^{(2)}(\beta). \tag{4.10}$$

In the proof of the results, we repeatedly apply the uniform convergence theorem of U-process (Nolan and Pollard (1987)). The imposed assumptions together with Lemmas 2.12-2.14 of Pakes and Pollard (1989) and Lemma 22 of Nolan and Pollard (1987) ensure that the underlying classes of functions are Euclidean.

Theorem 4.1. Under the validity of (4.1) and totally independent censoring, $\widetilde{\beta}_t \xrightarrow{p} \beta_t$ and $\sqrt{n}(\widetilde{\beta}_t - \beta_t) \xrightarrow{d} N_{p-1}(0, \Sigma_{0t}^{(1)})$ for $t \in (0, \tau_1]$ and τ_1 satisfying $P(X > \tau_1) > 0$, where $\Sigma_t^{(1)} = \{V_t^{(1)}\}^{-1} Cov(W_{t1}^{(1)}) \{V_t^{(1)}\}^{-1}$,

$$V_t^{(1)} = \frac{2\partial^2}{\partial\beta\partial\beta^T} h_t^{(1)}(\beta_t), W_{ti}^{(1)} = \frac{\partial}{\partial\beta} E[h_{ti1}^{(1)}(\beta_t) + h_{t1i}^{(1)}(\beta_t) + B_{t12i}(\beta_t) | X_i, Y_i, \delta_i]$$

$$B_{tijk}(\beta) = h_{tij}^{(1)}(\beta) \int_0^{X_i} \frac{dM_{Ck}(v)}{S_X(v)}, \text{ and } M_{Ci}(t) = I(X_i \le t)(1 - \delta_i) + \ln S_C(t \wedge X_i).$$

Proof. Define $H_t^{(1)}(\beta) = (n)_2^{-1} \sum_{i \neq j} h_{tij}^{(1)}(\beta)$, where

$$h_{tij}^{(1)}(\beta) = \frac{\delta_i}{S_C(X_i)} I(X_i \le t, X_j > t, \beta^T Y_i > \beta^T Y_j).$$

It is straightforward to show that β_t is the unique maximizer of $h_t^{(1)}(\beta) = E[h_{t12}^{(1)}(\beta)]$ and $h_t^{(1)}(\beta)$ is proportional to $\theta_t(\beta) S_C(t) S_T(t)(1 - S_T(t))$. From the uniform convergence of $\widetilde{S}_C(t)$ and U-process $H_t^{(1)}(\beta)$, one has

$$\sup_{\beta} |\widetilde{H}_t^{(1)}(\beta) - H_t^{(1)}(\beta)| = o_p(1) \text{ and } \sup_{\beta} |H_t^{(1)}(\beta) - h_t^{(1)}(\beta)| = o_p(1). \tag{4.11}$$

Following the argument of Newey and McFadden (1994), the consistency of $\widetilde{\beta}_t$ can be derived from $\sup_\beta |\widetilde{H}_t^{(1)}(\beta) - h_t^{(1)}(\beta)| = o_p(1)$, which is a direct consequence of (4.11).

For the asymptotic normality of $\widetilde{\beta}_t$, we first find an appropriate quadratic approximation of $\widetilde{H}_t^{(1)}(\beta)$ around $o_p(1)$ neighborhoods of β_t. The Taylor expansion of

$(\widetilde{H}_t^{(1)}(\beta)-\widetilde{H}_t^{(1)}(\beta_t))$ at $\widetilde{S}_C(t)=S_C(t)$ and the martingale representation of $\sqrt{n}(\widetilde{S}_C(t)-S_C(t))$ in (2.18) imply

$$(\widetilde{H}_t^{(1)}(\beta)-\widetilde{H}_t^{(1)}(\beta_t))=\frac{1}{(n)_3}\sum_{i\neq j\neq k}\{h_{tij}^{(1)}(\beta)-h_{tij}^{(1)}(\beta_t)+B_{tijk}(\beta)-B_{tijk}(\beta_t)\}+o_p(n^{-1}) \tag{4.12}$$

for β satisfying $||\beta-\beta_t||=o_p(1)$. Along the same lines as the proof of Theorem 4 in Sherman (1993) and the property $E[B_{tijk}(\beta)|X_s,Y_s,\delta_s]=0, s\in\{i,j\}$, for β satisfying $||\beta-\beta_t||=o_p(1)$, (4.12) can be further approximated by

$$\frac{1}{2}(\beta-\beta_t)^TV_t^{(1)}(\beta-\beta_t)+(\frac{1}{n}\sum_{i=1}^{n}W_{ti}^{(1)})^T(\beta-\beta_t)+o_p(||\beta-\beta_t||^2)+o_p(n^{-1}). \tag{4.13}$$

Together with Theorems 1-2 of Sherman (1993), $\sqrt{n}(\widetilde{\beta}_t-\beta_t)=n^{-1}\sum_{i=1}\{V_t^{(1)}\}^{-1}W_{ti}^{(1)}+o_p(1)$ and converges to a $(p-1)$-variate normal distribution with mean zero and variance-covariance matrix $\Sigma_t^{(1)}$. □

The large sample properties of $(\widetilde{TPR}_t(c,\widetilde{\beta}_t),\widetilde{FPR}_t(c,\widetilde{\beta}_t))$ and $\widetilde{\theta}_t(\widetilde{\beta}_t)$ are given in the following theorem.

Theorem 4.2. Suppose that the conditions in Theorem 4.1 are satisfied. For $t\in(0,\tau_1]$, $\sqrt{n}(\widetilde{TPR}_t(c,\widetilde{\beta}_t)-TPR_t(c,\beta_t),\widetilde{FPR}_t(c,\widetilde{\beta}_t)-FPR_t(c,\beta_t))^T$ converges to a bivariate Gaussian process in c with mean zero and variance-covariance matrix $\Sigma_{1t}^{(1)}(c)$, where $\Sigma_{1t}^{(1)}(c)$ is

$$Cov\left((A_{ti}^{(1)}(c,\beta_t)+\frac{\partial TPR_t(c,\beta_t)}{\partial\beta}(V_t^{(1)})^{-1}W_{ti}^{(1)},A_{ti}^{*(1)}(c,\beta_t)+\frac{\partial FPR_t(c,\beta_t)}{\partial\beta}(V_t^{(1)})^{-1}W_{ti}^{(1)})^T\right)$$

with

$$A_{ti}^{(1)}(c,\beta)=\phi_{ti}(c,\beta)-TPR_t(c,\beta)+E[\phi_{tj}(c,\beta)(\int_0^{X_j}\frac{dM_{Ci}(u)}{S_X(u)}-\int_0^t\frac{dM_{Ti}(u)}{S_X(u)})|X_i,\delta_i,Y_i],$$

$$A_{ti}^{*(1)}(c,\beta)=\frac{I(X_i>t,\beta^TY_i>c)}{S_X(t)}-FPR_t(c,\beta)-\frac{FPR_t(c,\beta)}{S_X(t)}(I(X_i>t)-S_X(t)),$$

$$\phi_{ti}(c,\theta_t) = \frac{\delta_i I(\beta_t^T Y_i > c)}{S_C(X_i)(1-S_T(t))} \text{ and } M_{Ti}(u) = I(X_i \le u)\delta_i + \ln S_T(t \wedge X_i).$$

Moreover, $\sqrt{n}(\widetilde{\theta}_t(\widetilde{\beta}_t) - \theta_t(\beta_t))$ converges to a normal distribution with mean zero and variance $Var(v_{it}^{(1)})$, where

$$\begin{aligned} v_{it}^{(1)} &= \frac{E[h_{tij}^{(1)}(\beta_t) + h_{tji}^{(1)}(\beta_t) + B_{tjki}(\beta_t)|X_i,\delta_i,Y_i] - 2h_t^{(1)}(\beta_t)}{S_C(t)S_T(t)(1-S_T(t))} \\ &+ \frac{h_t^{(1)}(\beta_t)}{S_C(t)S_T(t)(1-S_T(t))}\left(\int_0^t \frac{dM_{Ci}(u)}{S_X(u)} + \frac{1-2S_T(t)}{1-S_T(t)}\int_0^t \frac{dM_{Ti}(u)}{S_X(u)}\right). \end{aligned}$$

Proof. Let

$$Z_{\widetilde{TPR}_t}(c,\beta) = \sqrt{n}(\widetilde{TPR}_t(c,\beta) - TPR_t(c,\beta)).$$

By the first order Taylor expansion of $Z_{\widetilde{TPR}_t}(c,\beta)$ at $(\widehat{S}_C(X_i), \widehat{S}_T(t)) = (S_C(X_i), S_T(t))$ and the uniform convergence of $\widehat{S}_C(t)$ and $\widehat{S}_T(t)$, $Z_{\widetilde{TPR}_t}(c,\beta)$ can be uniformly approximated by

$$\frac{\sqrt{n}}{(n)_2}\sum_{i\neq j}\phi_{ti}(c,\beta) - TPR_t(c,\beta) - \phi_{ti}(c,\beta)\left(\frac{\widehat{S}_C(X_i) - S_C(X_i)}{\widehat{S}_C(X_i)} + \frac{\widehat{S}_T(t) - S_T(t)}{1-S_T(t)}\right). \quad (4.14)$$

Together with the martingale representations of $\widehat{S}_C(t)$ and $\widehat{S}_T(t)$, the term in (4.14) is asymptotically uniformly equivalent to

$$\frac{\sqrt{n}}{(n)_2}\sum_{i\neq j}\phi_{ti}(c,\beta) - TPR_t(c,\beta) + \phi_{ti}(c,\beta)\left(\int_0^{X_i}\frac{dM_{Cj}(u)}{S_X(u)} - \int_0^t \frac{dM_{Tj}(u)}{S_X(u)}\right) \quad (4.15)$$

over all (c,β). By the decomposition of a U-statistic into the sum of degenerate U-statistics and Corollary 4 of Sherman (1994a), the term in (4.15) can be approximated by $n^{-1/2}\sum_{i=1}^n A_{ti}^{(1)}(c,\beta)$. The functional central limit theorem further ensures that $Z_{\widetilde{TPR}_t}(c,\beta)$ converges weakly to Gaussian process in (c,β). From Theorem 4.1 and the equicontinuity of $Z_{\widetilde{TPR}_t}(c,\beta)$ in (c,β), it follows that $\sqrt{n}(\widetilde{TPR}_t(c,\widetilde{\beta}_t) - TPR_t(c,\beta_t))$ is uniformly equivalent to

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (A_{ti}^{(1)}(c,\beta_t) + \frac{\partial TPR_t(c,\beta_t)}{\partial\beta}(V_t^{(1)})^{-1}W_{ti}^{(1)}). \quad (4.16)$$

Similar to the derivation for (4.16), $\sqrt{n}(\widetilde{FPR}_t(c,\widetilde{\beta}_t) - FPR_t(c,\beta_t))$ have the following asymptotic representation

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(A_{ti}^{*(1)}(c,\beta_t) + \frac{\partial FPR_t(c,\beta_t)}{\partial\beta}(V_t^{(1)})^{-1}W_{ti}^{(1)}). \tag{4.17}$$

Thus, we conclude form (4.16)-(4.17) that $\sqrt{n}(\widetilde{TPR}_t(c,\widetilde{\beta}_t) - TPR_t(c,\beta_t), \widetilde{FPR}_t(c,\widetilde{\beta}_t) - FPR_t(c,\beta_t))^T$ converges weakly to a bivariate Gaussian process in c.

From (4.13) and $||\widetilde{\beta}_t - \beta_t|| = O_p(n^{-1/2})$, one has

$$\sqrt{n}(\widetilde{H}_t^{(1)}(\widetilde{\beta}_t) - h_t^{(1)}(\beta_t)) = \sqrt{n}(\widetilde{H}_t^{(1)}(\beta_t) - h_t^{(1)}(\beta_t)) + o_p(1) \tag{4.18}$$

It follows from (4.18) that $\sqrt{n}(\widetilde{\theta}_t(\widetilde{\beta}_t) - \theta_t(\beta_t))$ can be approximated by

$$\frac{\sqrt{n}}{(n)_3}\sum_{i\neq j\neq k}\frac{h_{tij}^{(1)}(\beta_t) + B_{tijk}(\beta_t) - h_t^{(1)}(\beta_t)(1 - \int_0^t \frac{dM_{Ci}(u)}{S_X(u)} - \frac{1-2S_T(t)}{1-S_T(t)}\int_0^t \frac{dM_{Ti}(u)}{S_X(u)})}{S_C(t)S_T(t)(1-S_T(t))}. \tag{4.19}$$

Since (4.19) is a mean zero U-statistics and can be approximated by $n^{-1/2}\sum_{i=1}^n v_{it}^{(1)}$, the proof is completed by applying the central limit theorem. □

For the asymptotic properties of $\widehat{\beta}_t$, $(\widehat{TPR}_t(c,\widehat{\beta}_t), \widehat{FPR}_t(c,\widehat{\beta}_t))$, and $\widehat{\theta}_t(\widehat{\beta}_t)$, the following conditions are made for Theorems 4.3-4.4 and the technical lemma.

(A1) $n\lambda_M^{2p+1}(\ln n)^{-1} < \infty$ and $n\lambda_m^{2p} \to \infty$, where $\lambda_m = \min\{\lambda_1, \cdots, \lambda_p\}$ and $\lambda_M = \max\{\lambda_1, \cdots, \lambda_p\}$.

(A2) $K(\cdot)$ is a $(p+1)$th order kernel function with bounded variation.

(A3) $g_\beta(y) = P(\beta^T Y > \beta^T y)$ has a uniformly bounded second derivative with respect to β.

(A4) The $(p+1)$th order derivatives of $\xi_y(x,\delta,t)$ in Lemma A and $\frac{\partial}{\partial\beta}g_{\beta_t}(y)$ with respect to y are uniformly bounded.

(A5) $P(\beta^T(Y_i - Y_j) > 0, \beta_t^T(Y_i - Y_j) \leq 0) \leq M||\beta - \beta_t||$ for some positive bounded constant M.

Theorem 4.3. Suppose that assumptions (A1)-(A5) in the Appendix are satisfied. Under the validity of model in (4.1) and marker-dependent censoring, $\widehat{\beta}_t \xrightarrow{p} \beta_t$ and $\sqrt{n}(\widehat{\beta}_t - \beta_t) \xrightarrow{d} N_{p-1}(0, \Sigma_t^{(2)})$ for $t \in (0, \tau_2]$ and τ_2 satisfying $\inf_y S_X(\tau_2, y) > 0$, where $\Sigma_t^{(2)} = \{V_t^{(2)}\}^{-1} Cov(\bar{W}_{ti}^{(2)})\{V_t^{(2)}\}^{-1}$,

$$V_t^{(2)} = \frac{2\partial^2}{\partial\beta\partial\beta^T} E[h_{t12}^{(2)}(\beta_t)], \bar{W}_{ti}^{(2)} = W_{ti}^{(2)} + W_{ti}^{*(2)},$$

$$W_{ti}^{(2)} = \frac{\partial}{\partial\beta} E[h_{tij}^{(2)}(\beta_t) + h_{tji}^{(2)}(\beta_t)|X_i, \delta_i, Y_i], \text{ and } W_{ti}^{*(2)} = \xi_{Y_i}(X_i, \delta_i, t) f_Y(Y_i) \frac{\partial}{\partial\beta} g_{\beta_t}(Y_i).$$

Proof. Let $H_t^{(2)}(\beta) = (n)_2^{-1} \sum_{i \neq j} h_{tij}^{(2)}(\beta)$ with $h_{tij}^{(2)}(\beta) = S_T(t|Y_j) I(\beta^T Y_i > \beta^T Y_j)$. Similar to the proof of Theorem 4.1, the consistency of $\widehat{\beta}_t$ is established when

$$\sup_\beta |\widehat{H}_t^{(2)}(\beta) - h_t^{(2)}(\beta)| = o_p(1)$$

holds. Here, $h_t^{(2)}(\beta) = E[h_{tij}^{(2)}(\beta)] = \theta_t(\beta) S_T(t)(1 - S_T(t)) + 0.5 S_T^2(t)$ has a unique maximizer β_t. From the boundedness of indicator functions, the above property can be obtained through

$$\sup_y |\widehat{S}_T(t|y) - S_T(t|y)| = o_p(1) \text{ and } \sup_\beta |H_t^{(2)}(\beta) - h_t^{(2)}(\beta)| = o_p(1), \tag{4.20}$$

which is a direct consequence of Corollary 2.1 of Dabrowska (1989) and the uniform convergence of U-process $H_t^{(2)}(\beta)$ to $h_t^{(2)}(\beta)$.

From the consistency of $\widehat{\beta}_t$ to β_t, one can further restrict β in $o_p(1)$ neighborhoods of β_t in the following derivation. For β satisfying $||\beta - \beta_t|| = o_p(1)$, it is implied from Lemma A below, (A1), and (A5) that

$$\widehat{H}_t^{(2)}(\beta) - \widehat{H}_t^{(2)}(\beta_t) = (H_t^{(2)}(\beta) - H_t^{(2)}(\beta_t)) + (U(\beta) - U(\beta_t)) + o_p(\frac{||\beta - \beta_t||}{n^{1/2}}) + o_p(n^{-1}), \tag{4.21}$$

where $U(\beta) = (n)_3^{-1} \sum_{i \neq j \neq k} f_{tk}^{(\lambda)}(Y_j) I(\beta^T Y_i > \beta^T Y_j)$. Using the same argument for (4.13), $(H_t^{(2)}(\beta) - H_t^{(2)}(\beta_t))$ can be re-expressed as

$$\frac{1}{2}(\beta - \beta_t)^T V_t^{(2)}(\beta - \beta_t) + (\frac{1}{n}\sum_{i=1}^n W_{ti}^{(2)})^T(\beta - \beta_t) + o_p(||\beta - \beta_t||^2) + o_p(n^{-1}) \tag{4.22}$$

uniformly over $o_p(1)$ neighborhoods of β_t. As for the term $U(\beta)$ in (4.21), there exist degenerate U-statistics $U_k(\beta)$, $k = 1, 2, 3$, of order k so that

$$U(\beta) - U(\beta_t) = \sum_{k=1}^{3} (U_k(\beta) - U_k(\beta_t)). \tag{4.23}$$

The property $U_1(\beta) = n^{-1} \sum_{i=1}^{n} E[f_{ti}^{(\lambda)}(Y_1) g_\beta(Y_1) | X_i, Y_i, \delta_i]$, assumptions (A1)-(A4), and the Taylor expansion theorem ensure that

$$U_1(\beta) - U_1(\beta_t) = \{\frac{1}{n} \sum_{i=1}^{n} W_{ti}^{*(2)}\}^T (\beta - \beta_t) + o_p(||\beta - \beta_t||^2) + o_p(\frac{||\beta - \beta_t||}{n^{1/2}}) \tag{4.24}$$

uniformly over $o_p(1)$ neighborhood of β_t. Note that $U_{2ij}(\beta)$ and $U_{3ijk}(\beta)$ are the kernel functions of $U_2(\beta)$ and $U_3(\beta)$ and satisfy $\lambda_m^p |U_{2ij}(\beta) - U_{2ij}(\beta_t)| \leq M||\beta - \beta_t||$ and $\lambda_m^p |U_{3ijk}(\beta) - U_{3ijk}(\beta_t)| \leq M$. It is implied from Theorem 3 of Sherman (1994b) and assumption (A1) that for β satisfying $||\beta - \beta_t|| = o_p(1)$,

$$|U_2(\beta) - U_2(\beta_t)| = o_p(\frac{||\beta - \beta_t||}{n^{1/2}}) \text{ and } |U_3(\beta) - U_3(\beta_t)| = o_p(n^{-1}). \tag{4.25}$$

From (4.21)-(4.25) and Theorem 1 of Sherman (1993), $||\widehat{\beta}_t - \beta_t|| = O_p(n^{-1/2})$ is derived and, hence, $(\widehat{H}_t^{(2)}(\widehat{\beta}_t) - \widehat{H}_t^{(2)}(\beta))$ is derived to be

$$\frac{1}{2}(\widehat{\beta}_t - \beta_t)^T V_t^{(2)} (\widehat{\beta}_t - \beta_t) + \{\frac{1}{n} \sum_{i=1}^{n} \bar{W}_{ti}^{(2)}\}^T (\widehat{\beta}_t - \beta_t) + o_p(n^{-1}). \tag{4.26}$$

By applying Theorem 2 of Sherman (1993), the asymptotic representation $\sqrt{n}(\widehat{\beta}_t - \beta_t) = n^{-1} \sum_{i=1}^{n} \{V_t^{(2)}\}^{-1} \bar{W}_{ti}^{(2)} + o_p(1)$ can be obtained and is shown to converge to a $(p-1)$-variate normal distribution with mean zero and variance-covariance matrix $\Sigma_t^{(2)}$. □

Theorem 4.4. Suppose that the conditions in Theorem 4.3 are satisfied. For $t \in (0, \tau_2]$, $\sqrt{n}(\widehat{TPR}_t(c, \widehat{\beta}_t) - TPR_t(c, \beta_t), \widehat{FPR}_t(c, \widehat{\beta}_t) - FPR_t(c, \beta_t))^T$ converges to a bivariate Gaussian process in c with mean zero and variance-covariance matrix $\Sigma_{1t}^{(2)}(c)$, where $\Sigma_{1t}^{(2)}(c)$ is

$$Cov\left((A_{ti}^{*(2)}(c, \beta_t) + \frac{\partial TPR_t(c, \beta_t)}{\partial \beta} (V_t^{(2)})^{-1} \bar{W}_{ti}^{(2)}, A_{ti}^{(2)}(c, \beta_t) + \frac{\partial FPR_t(c, \theta_{0t})}{\partial \beta} (V_t^{(2)})^{-1} \bar{W}_{ti}^{(2)})^T \right)$$

with

$$A_{ti}^{(2)}(c,\beta) = \frac{(I(\beta^T Y_j > c) - FPR_t(c,\beta))(\xi_{Y_i}(X_i,\delta_i,t)f_Y(Y_i) + S_T(t|Y_i))}{S_T(t)} \text{ and}$$

$$A_{ti}^{*(2)}(c,\beta) = \frac{(I(\beta^T Y_i > c) - TPR_t(c,\beta))((1 - S_T(t|Y_i)) - \xi_{Y_i}(X_i,\delta_i,t)f_Y(Y_i))}{1 - S_T(t)}.$$

Moreover, $\sqrt{n}(\widehat{\theta}_t(\widehat{\beta}_t) - \theta_t(\beta_t))$ converges to a normal distribution with mean zero and variance $Var(v_{it}^{(2)})$, where

$$\begin{aligned} v_{it}^{(2)} &= \frac{E[h_{tij}^{(2)}(\beta_t) + h_{tji}^{(2)}(\beta_t)|X_i,\delta_i,Y_i] + g_{\beta_t}(Y_i)\xi_{Y_i}(X_i,\delta_i,t)f_Y(Y_i) - 2h_t^{(2)}(\beta_t)}{S_T(t)(1 - S_T(t))} \\ &- \frac{S_T(t|Y_i) + \xi_{Y_i}(X_i,\delta_i,t)f_Y(Y_i) - S_T(t)}{2(1 - S_T(t))^2}. \end{aligned}$$

Proof. Let

$$Z_{\widehat{FPR}_t}(c,\beta) = \sqrt{n}(\widehat{FPR}_t(c,\beta) - FPR_t(c,\beta)).$$

From Lemma A, $Z_{\widehat{FPR}_t}(c,\beta)$ has the following asymptotic representation

$$\frac{\sqrt{n}}{(n)_2}\sum_{i\neq j}\frac{(I(\beta^T Y_i > c) - FPR_t(c,\beta))(S_T(t|Y_i) + f_{tj}^{(\varsigma)}(Y_i))}{S_T(t)}. \tag{4.27}$$

By the decomposition of a U-statistic into the sum of degenerate U-statistics and Corollary 4 of Sherman (1994a), the term in (4.27) can be further approximated by $n^{-1/2}\sum_{i=1}^n A_{ti}^{(2)}(c,\beta)$. The functional central limit theorem ensures that $Z_{\widehat{FPR}_t}(c,\beta)$ converges weakly to a Gaussian process in (c,β). From Theorems 4.3 and the equicontinuity of $Z_{\widehat{FPR}_t}(c,\beta)$ in (c,θ_t), $\sqrt{n}(\widehat{FPR}_t(c,\widehat{\beta}_t) - FPR_t(c,\beta_t))$ is shown to be uniformly equivalent to

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (A_{ti}^{(2)}(c,\beta_t) + \frac{\partial FPR_t(c,\beta_t)}{\partial\beta}(V_t^{(2)})^{-1}\bar{W}_{ti}^{(2)}). \tag{4.28}$$

Similar to the derivation for (4.28), $\sqrt{n}(\widehat{TPR}_t(c,\widehat{\beta}_t) - TPR_t(c,\beta_t))$ has the following asymptotic representation

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (A_{ti}^{*(2)}(c,\beta_t) + \frac{\partial TPR_t(c,\beta_t)}{\partial\beta}(V_t^{(2)})^{-1}\bar{W}_{ti}^{(2)}). \tag{4.29}$$

Form (4.29), it can be shown that $\sqrt{n}(\widehat{TPR_t}(c,\widehat{\beta}_t)-TPR_t(c,\beta_t), \widehat{FPR_t}(c,\widehat{\beta}_t)-FPR_t(c,\beta_t))^T$ converges weakly to bivariate Gaussian processes in c.

From (4.26) and $||\widehat{\beta}_t - \beta_t|| = O_p(n^{-1/2})$, one has

$$\sqrt{n}(\widehat{H}_t^{(2)}(\widehat{\beta}_t) - h_t^{(2)}(\beta_t)) = \sqrt{n}(\widehat{H}_t^{(2)}(\beta_t) - h_t^{(2)}(\beta_t)) + o_p(1). \tag{4.30}$$

By Lemma A and (4.30), the asymptotic representation of $\sqrt{n}(\widehat{\theta}_t(\widehat{\beta}_t) - \theta_t(\beta_t))$ is derived as

$$\frac{\sqrt{n}}{(n)_3}\sum_{i\neq j\neq k}\left(\frac{h_{tij}^{(2)}(\beta_t)(1+f_{tk}^{(\varsigma)}(Y_j)) - h_t^{(2)}(\beta_t)}{S_T(t)(1-S_T(t))} - \frac{S_T(t|Y_i)(1+f_{tj}^{(\varsigma)}(Y_i)) - S_T(t)}{2(1-S_T(t))^2}\right). \tag{4.31}$$

Since the random quantity in (4.31) is a mean zero U-statistics, it can be approximated by $n^{-1/2}\sum_{i=1}^n v_{it}^{(2)}$ and, hence, the theorem is obtained. □

Since the asymptotic variance-covariance matrices of $\widetilde{\beta}_t$, $\widetilde{TPR_t}(c,\widetilde{\beta}_t)$, $\widetilde{FPR_t}(c,\widetilde{\beta}_t)$, $\widehat{\beta}_t$, $\widehat{TPR_t}(c,\widehat{\beta}_t)$, and $\widehat{FPR_t}(c,\widehat{\beta}_t)$ involve the derivatives of unknown quantities, which are complicated and hard to estimate directly under the nonparametric setting, the bootstrap variance estimates are considered in our numerical studies and applications. It is found from (4.18) and (4.30) that $\widehat{\beta}_t$ and $\widetilde{\beta}_t$ have no effect on the asymptotic variances of the corresponding classification accuracies $\widehat{\theta}_t(\widehat{\beta}_t)$ and $\widetilde{\theta}_t(\widetilde{\beta}_t)$. That is, the performance of $\widehat{\beta}_t^T Y$ or $\widetilde{\beta}_t^T Y$ are asymptotically equivalent to the "observed" true linear predictors. Thus, the statistical inferences for $\theta_t(\beta_t)$ can be made through the methods developed in Chapter 2 directly.

Lemma A. Suppose that assumptions (A1)-(A2) are satisfied. For $t \in (0, \tau_2]$,

$$\widehat{S}_T(t|y) - S_T(t|y) = \frac{1}{n}\sum_{i=1}^n f_{ti}^{(\lambda)}(y) + r_n(t,y), \tag{4.32}$$

where $f_{ti}^{(\lambda)}(y) = \xi_y(X_i,\delta_i,t)K_\lambda(Y_i - y) - E[\xi_y(X_i,\delta_i,t)K_\lambda(Y_i - y)]$ and $\sup_y |r_n(t,y)| = O_p(\ln n/(n\lambda_M^p))$ with $\xi_y(X_i,\delta_i,t) = -S_T(t|y)\int_0^t d_u M_i(u,y)/S_X(u,y)$ and $M_i(t,y) = I(X_i \le t)\delta_i + \ln S_T(t\wedge X_i|y)$.

Proof. Paralleling the proof of Theorem 3.2 in Du and Akritas (2002), one can derive that $\widehat{S}_T(t|y) - S_T(t|y) = n^{-1}\sum_{i=1}^{n} \xi_y(X_i, \delta_i, t) K_\lambda(Y_i - y) + r_n^*(t, y)$ with $\sup_y |r_n^*(t, y)| = O_p(\ln n/(n\lambda_M^p))$. By assumption (A1) and the property $E[\xi_{Y_i}(X_i, \delta_i, t)|Y_i] = 0$, we have $\sup_y |E[\xi_{Y_j}(X_i, \delta_i, t) K_\lambda(Y_i - y)]| = O(\lambda_M^{p+1})$ and, hence, (4.32). □

Chapter 5

Simulations and Applications

We conduct a series of Monte Carlo simulations to investigate the finite sample properties of the proposed estimators and inference procedures. Two empirical examples form the Angiography CAD Study and the ACTG 175 Study are analyzed to demonstrate the usefulness of our methodologies.

5.1 Monte Carlo Simulations

In the succeeding numerical studies, data are repeatedly generated 500 times in each simulation setting. The performances of our methods are evaluated under different sample sizes ($n = 250$ and 500), censoring rates ($c.r. = 30\%$ and 50%), and censoring mechanisms over various time points t_q's and time intervals. In the exhibited tables, t_q is used to denote the qth quantile of the failure time T.

5.1.1 Scenario I - Time-Dependent AUC and PAUC

For the time-dependent AUC and PAUC, the continuous biomarker Y is generated from a standard normal distribution. Conditioning on $Y = y$, T and C are independently generated from a lognormal distribution with parameters $\mu = -0.22y$ and $\sigma = 0.3$, and an exponential distribution with scale parameter $b\{2I(y < 0) + I(y \geq 0)\}$.

Moreover, under totally independent censoring, C is independently generated from another exponential distribution with scale parameter b. The constant b is mainly used to control the expected censoring rates. With α =0.1, 0.3, and 1, the estimators $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ and $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ are evaluated at the selected time points $t_{0.4}$, $t_{0.5}$, and $t_{0.6}$. Moreover, the simultaneous confidence bands for $\theta_t(q_{\alpha t})$ over the subintervals $[t_{0.4}, t_{0.5}]$ and $[t_{0.4}, t_{0.6}]$ of interest are constructed. Since a small portion of cases or controls occur outside $[t_{0.4}, t_{0.6}]$, the simulation results are only presented within this time period.

Since the smoothing estimator $\widehat{S}(t|y)$ is involved in the estimator $\widehat{\theta}_t(\widehat{q}_{\alpha t})$, an appropriate smoothing parameter becomes necessary. It usually attempts to select a bandwidth that minimizes the asymptotic mean squared error of an estimator, which is obtained by using the plug-in method for unknown parameters. This approach, however, would lead to further bandwidth selection problems and is infeasible in our current setting. Here, we propose a simple and easily implemented data-driven method to find a bandwidth λ_{opt} which is the minimizer of the integrated squared error:

$$ISE(\lambda) = \int_0^1 \{\widehat{S}_e(u) - (1-u)\}^2 dN_{ei}(u), \tag{5.1}$$

where $\widehat{S}_e(u)$ is the Kaplan-Meier estimator computed based on the data $\{e_i, \delta_i\}_{i=1}^n$, $e_i = 1 - \widehat{S}_T^{(-i)}(X_i|Y_i)$, $\widehat{S}_T^{(-i)}(t|y)$ is obtained as $\widehat{S}_T(t|y)$ with the ith observation (X_i, δ_i, Y_i) being deleted, and $N_{ei}(u) = \delta_i I(e_i \leq u)$. The rationale behind (5.1) is that $\{1 - S_T(X_i|Y_i), \delta_i\}_{i=1}^n$ can be shown to be an independent censored sample from a standard uniform distribution under the validity of marker-dependent censoring. To assess the performance of (5.1), the estimators $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ and $\widehat{\Psi}_{\alpha i}(t)$'s are computed using the cross-validation bandwidth λ_{opt} and the subjective ones of 0.01 and 0.2.

Tables 5.1-5.4 summarize the simulation results under the setting of totally independent censoring. It can be found that $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ and $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ with $\lambda = \lambda_{opt}$ have relatively small biases especially for small α. These tables also show that the standard deviations of both $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ and $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ are are very close to the averages of $\widetilde{\Sigma}_\alpha^*(t,t)$ and $\widehat{\Sigma}_\alpha(t,t)$. The standard deviation of $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ is found to be sightly smaller than that of $\widehat{\theta}_t(\widehat{q}_{\alpha t})$. Interestingly, the performance of $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ is stable with high censoring

rate while the standard deviation becomes larger. Compare with $\widehat{\theta}_t(\widehat{q}_{\alpha t})$, the bias of $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ is more serious to heavy censoring. The empirical pointwise and simultaneous coverage probabilities of $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ are further shown to be closer to the nominal level 0.95 than those of $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ especially for high censoring rate, and small quantile value and sample size. Overall, the performance of these methods can be significantly improved by increasing sample size. As for the smoothing estimators for $\theta_t(q_{\alpha t})$ with $\lambda = 0.01$ and 0.2, the poor estimates are observed.

For the simulation setting of marker-dependent censoring, the related summary statistics and empirical coverage probabilities are exhibited in tables 5.5-5.8. A similar conclusion under totally independent censoring can be drawn for $\widehat{\theta}_t(\widehat{q}_{\alpha t})$. As for the estimator $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$, an apparent bias arises especially for small α. The biases of $\widehat{\theta}_t(\widehat{q}_{0.1t})$ and $\widetilde{\theta}_t(\widetilde{q}_{0.1t})$ appear to be larger because they are computed based on small proportion of subjects in control group. Furthermore, the coverage probabilities of $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ systematically deviate from 0.95 even with large sample size. Although $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ is robust to violation of marker-dependent censoring, the performance of $\widehat{\theta}_t(\widehat{q}_{\alpha t})$ in our simulation setting is detected to be better. Generally, the proposed bandwidth selection procedure provides satisfactory results.

5.1.2 Scenario Ⅱ - Time-Dependent AUC Regression Model

A univariate covariate Z is set to follow an exponential distribution with scale parameter 0.5. Conditioning on Z and a Bernoulli latent variable v with parameter 0.5, T and Y are separately generated from exponential distributions with scale parameters 3^{-v} and Z^{2v}. The censoring time C is designed to be an exponential distribution with different scale parameters, which result in the expected censoring rates. The Kaplan-Meier estimator $\widetilde{S}_C(t)$ is further used to estimate $S_C(t|z^*)$ and $G_i(t, z^*) = -S_C(t)\int_0^t dM_{Ci}(u)/S_X(u)$ is used to estimate $\widehat{G}_i(t, z^*) = -\widetilde{S}_C(t)\int_0^t d\widetilde{M}_{Ci}(u)/\widetilde{S}_X(u)$. Under the above setting, it can be derived that the true model for $\theta_t(Z_i, Z_j)$ has the linear form

$$\theta_t^{(T)}(Z_i, Z_j) = \gamma_{t1}Z_{ij1} + \gamma_{t2}Z_{ij2} + \gamma_{t3}Z_{ij3} + \gamma_{t4}Z_{ij4},$$

where $Z_{ij1} = 1$, $Z_{ij2} = (1 + Z_i^{-2})^{-1}$, $Z_{ij3} = (1 + Z_j^2)^{-1}$, and $Z_{ij4} = Z_i^2(Z_i^2 + Z_j^2)^{-1}$ with the corresponding coefficients $\gamma_{t1} = 0.125(e^{-t} - e^{-2t})$, $\gamma_{t2} = 0.25(e^{-t} - e^{-4t})$, $\gamma_{t3} = 0.25(e^{-3t} - e^{-4t})$, and $\gamma_{t4} = 0.25(e^{-3t} - e^{-6t})$. In the simulation study, the true response function $h^{(T)}(u) = u$ with the designed covariates Z_{ijk}'s are used in model (3.1) to estimate γ_t. Moreover, the logistic regression model $h^{(W)}(u) = e^u/(1 + e^u)$ is considered as a working model. Both the true and working models are used to estimate the scientific relevant measures $\theta_t^{(T)}(z_0, z_0)$, $z_0 = 1.5, 2.5$. We evaluate the finite sample properties of the proposed estimators and inference procedures at the selected time points $t_{0.3}, t_{0.5}$, and $t_{0.7}$.

It is indicated from table 5.9 that the estimators $\widehat{\gamma}_t$ behave well in the mid time period no matter how the sample size and the censoring rate change. However, apparent biases appear for times near the boundary due to the small sample size and the high censoring rate. This situation is greatly improved as the sample size increases. All the standard deviation is also found to be nicely estimated by our proposed empirical estimation methods. With moderate sample size and censoring rate, most of the empirical coverage probabilities of pointwise confidence intervals are close to the nominal level of 0.95. It is further shown in table 5.12 that the empirical coverage probabilities of the simultaneous confidence bands for the intreval $[t_{0.5}, t_{0.6}]$ are closer to the nominal level than those of $[t_{0.3}, t_{0.7}]$. As expected, the adequacy of the constructed confidence bands relies on the change of sample size and censoring rate. We further detect in tables 5.10-5.11 that the estimates $\widehat{\theta}^{(T)}(z_0, z_0)$'s and $\widehat{\theta}^{(W)}(z_0, z_0)$'s computed based on $h^{(T)}(u)$ and $h^{(W)}(u)$, respectively, are close to the corresponding true values $\theta^{(T)}(z_0, z_0)$'s. The strong performance of the time-varying coefficient logistic regression model results in wide usages in applications and is evidenced in our simulation study.

5.1.3 Scenario Ⅲ - Optimal Composite Markers

In this simulation study, two markers Y_1 and Y_2 are considered in the time-varying coefficient EGLM. We first generate $Z = (Z_1, Z_2)^T$ from a bivariate normal distribution with mean $(-0.5, -2)$, standard deviation $(0.22, 0.25)$, and correlation coefficient -0.7. Conditioning on Z, the failure time T is designed to follow a lognormal distribution with parameters $\mu(Z) = -(Z_1 + Z_2)/1.6$ and $\sigma^2(Z) = \exp(2Z_2)$. The markers Y_1 and Y_2 are specified to be $1/\sigma(Z)$ and $-\mu(Z)/\sigma(Z)$. Under the above design, the linear composite marker of the form $(Y_1 + \beta_t Y_2)$ with $\beta_t = 1/\ln t$ can be shown to be optimal at time t. Under totally independent censoring, C is generated from an exponential distribution with scale parameter a. With marker-dependent censoring, conditioning on $(Y_1, Y_2) = (y_1, y_2)$, C is generated from a gamma distribution with shape parameter (y_2/a) and scale parameter $(y_1^2/10)$. The constant a in both settings is used to obtained the censoring rates of 30% and 50%. The estimators and inference procedures are evaluated at selected time points of $t_{0.3}$, $t_{0.5}$, and $t_{0.7}$. In the estimation of $\widehat{\beta}_t$ and $\widehat{\theta}_t(\widehat{\beta}_t)$, the standard normal density function is used as a kernel function. For the smoothing parameters involved in $\widetilde{S}_T(t|y)$, an appropriate selection procedure becomes necessary for simplifying the process and preventing the investigators from arbitrarily choosing the bandwidth in practical implementation. One possible method, for instance, is to select small enough smoothing parameters so that $\widetilde{\beta}_t$ and $\widehat{\beta}_t$ are comparable. In the current simulation study, the bandwidths are all set to be 0.05.

Tables 5.13-5.14 exhibit the simulation results under totally independent censoring. One can see that the biases of both estimators $\widehat{\beta}_t$ and $\widetilde{\beta}_t$ are quite small, whereas the standard error of $\widetilde{\beta}_t$ is smaller than that of $\widehat{\beta}_t$. At the boundary time points, the standard errors of both estimates $\widehat{\beta}_t$ and $\widetilde{\beta}_t$ become substantially large especially when the sample size is small and the censoring rate is high. It can be observed that the bootstrap method provides good estimates of the standard deviations. The empirical coverage probabilities of the bootstrap confidence intervals are nearly the assigned

nominal level of 0.95 with moderate sample size and censoring rate. A similar conclusion about the biases of $\widehat{\theta}_t(\widehat{\beta}_t)$ and $\widetilde{\theta}_t(\widetilde{\beta}_t)$ can also be drawn. Here, the variation of $\widehat{\theta}_t(\widehat{\beta}_t)$ are detected to be slightly larger. The empirical coverage probabilities of the constructed confidence intervals based on the bootstrap analogue of $\widehat{\theta}_t(\widehat{\beta}_t)$ are more close to the nominal level than those based on the bootstrap analogue of $\widetilde{\theta}_t(\widetilde{\beta}_t)$. As one can expect, the performance of $\widetilde{\beta}_t$ is generally better than that of $\widehat{\beta}_t$ under totally independent censoring while both estimators $\widehat{\theta}_t(\widehat{\beta}_t)$ and $\widetilde{\theta}_t(\widetilde{\beta}_t)$ are comparable and acceptable.

Simulation results under marker-dependent censoring are provided in tables 5.15-5.16. One can see from these tables that the averages of $\widehat{\beta}_t$ and $\widehat{\theta}_t(\widehat{\beta}_t)$ are close to the corresponding true values at the selected time points. Interestingly, the bias of $\widetilde{\beta}_t$ is not apparent and shows the robustness against violation of marker-dependent censoring. The bootstrap standard errors and the empirical coverage probabilities are also close to the standard deviations of the estimators and the nominal level. On the other hand, the estimator $\widetilde{\theta}_t(\widetilde{\beta}_t)$ is very sensitive to the violation of totally independent censoring and is is evidenced by the occurrence of large biases. In summary, the performance of $\widehat{\theta}_t(\widehat{\beta}_t)$ is better under marker-dependent censoring, and $\widetilde{\beta}_t$ is also suggested to be used in the estimation due to its robustness. One advantage of applying $\widetilde{\beta}_t$ is that no complicated smoothing technique is involved.

5.2 Applications

5.2.1 ACTG 175 Study

We apply our methods to the data from the ACTG 175 study. Since patients with lower CD4 cell counts are more likely to have a higher risk of AIDS or death, a strictly decreasing transformation of CD4 cell counts, e.g. Y =-CD4, is used to ascertain that most of the AUC values fall within the range of 0.5 to 1. In our analysis, a class of appropriate survival models is applied to investigate the relationship between the censoring time and the CD4 cell counts, and no apparent association is detected.

We thus assume that totally independent censorship is reasonable and the estimator $\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$ is suggested to be used in this analysis. Due to a small number of failures and large variation of estimators at the initial weeks, we only provide the estimates within the time period after week 98. The nonparametric estimates of AUCs for all patients are displayed in figure 5.3 (a) together with the corresponding 0.95 pointwise and simultaneous confidence intervals, which are constructed from (2.25) and (2.28). A high classification accuracy curve of CD4 cell counts throughout appears within the time period of interest.

Let $Y^{(1)}$ and $Y^{(2)}$ be the CD4 cell biomarkers of non-therapy and therapy patients with the corresponding time-dependent PAUCs $\theta_t^{(1)}(q_{\alpha t}^{(1)})$ and $\theta_t^{(2)}(q_{\alpha t}^{(2)})$. Our aim in this data study is to evaluate the effect of prior therapy on the classification accuracy of CD4 cell counts. Currently, there is still no standard of clinically meaningful values of FPR for the PAUC in AIDS research, we consider the time-dependent PAUC with FPR less than 0.1, 0.3, and 1. Based on two independent data sets $\{X_i^{(1)}, \delta_i^{(1)}, Y_i^{(1)}\}_{i=1}^{n_1}$ and $\{X_i^{(2)}, \delta_i^{(2)}, Y_i^{(2)}\}_{i=1}^{n_2}$, the confidence intervals for $\theta_t^{(k)}(q_{\alpha t}^{(k)})$'s are constructed and compared with $\alpha^2/2$. It is detected from figure 5.1 that the time-dependent PAUCs $\theta_t^{(1)}(q_{\alpha t}^{(1)})$ and $\theta_t^{(2)}(q_{\alpha t}^{(2)})$ are significantly higher than $0.5\alpha^2$ for all α during the time period of interest. This indicates that CD4 cell counts are useful in classifying patient's survival time for all patients. For patients without prior therapy, a decreasing trend in $\theta_t^{(1)}(q_{\alpha t}^{(1)})$ over time is detected especially for small α. The decreasing trend is also found for $\theta_t^{(2)}(q_{\alpha t}^{(2)})$ before week 155, but it appears to increase after that time. For the time-dependent AUC, this classification measures stay very close to a constant throughout the study period for both groups. In view of the time-dependent AUC, the classification ability of CD4 cell counts seems to be irrelevant to the time parameter t.

The difference in the classification accuracies of $Y^{(1)}$ and $Y^{(2)}$ can be further measured by the summary index $\gamma_\alpha(t) = \theta_t^{(1)}(q_{\alpha t}^{(1)}) - \theta_t^{(2)}(q_{\alpha t}^{(2)})$. It is natural to estimate $\gamma_\alpha(t)$ by $\widetilde{\gamma}_\alpha(t) = \widetilde{\theta}_t^{(1)}(\widetilde{q}_{\alpha t}^{(1)}) - \widetilde{\theta}_t^{(2)}(\widetilde{q}_{\alpha t}^{(2)})$. Along the same lines as the proof in Section 2.2, we can derive that $\sqrt{n}(\widetilde{\gamma}_\alpha(t) - \gamma_\alpha(t))$ converges weakly to a mean zero Gaussian

process in t with variance-covariance function

$$\Gamma^*_\alpha(s,t) = \frac{1}{\kappa}E[\Psi^{*(1)}_{\alpha i}(s)\Psi^{*(1)}_{\alpha i}(t)] + \frac{1}{1-\kappa}E[\Psi^{*(2)}_{\alpha i}(s)\Psi^{*(2)}_{\alpha i}(t)] \tag{5.2}$$

provided that $n_1/n \to \kappa$ $(0 < \kappa < 1)$ as $n = (n_1 + n_2) \to \infty$, where $\Psi^{*(k)}_{\alpha i}(t)$ is a counterpart of $\Psi^*_{\alpha i}(t)$, $k = 1, 2$. To make inference on $\gamma_\alpha(t)$, $\Gamma^*_\alpha(s,t)$ is first estimated by

$$\widetilde{\Gamma}_\alpha(s,t) = \frac{n}{n_1^2}\sum_{i=1}^{n_1}\widetilde{\Psi}^{*(1)}_{\alpha i}(s)\widetilde{\Psi}^{*(1)}_{\alpha i}(t) + \frac{n}{n_2^2}\sum_{i=1}^{n_2}\widetilde{\Psi}^{*(2)}_{\alpha i}(s)\widetilde{\Psi}^{*(2)}_{\alpha i}(t). \tag{5.3}$$

A $(1-\varsigma)$ pointwise confidence interval for $\gamma_\alpha(t)$ and a $(1-\varsigma)$ simultaneous confidence band for $\{\gamma_\alpha(t) : t \in [\tau_1, \tau_2]\}$ are separately given via

$$\widetilde{\gamma}_\alpha(t) \pm Z_{\varsigma/2}\sqrt{\frac{\widetilde{\Gamma}^*_\alpha(t,t)}{n}} \text{ and } \{\widetilde{\gamma}_\alpha(t) \pm L^{(\gamma)}_\varsigma\sqrt{\frac{\widetilde{\Gamma}^*_\alpha(t,t)}{n}} : t \in [\tau_1, \tau_2]\} \tag{5.4}$$

with $L^{(\gamma)}_\varsigma$ being obtained as that in (2.4). It is revealed in figures 5.2 (a)-(c) that only $\gamma_\alpha(t)$ with α=0.1 and 0.3 tend to be positive before week 155 but the difference becomes negligible for all α's after that week. Thus, with small values of $FPR_t(y)$, a prior antiretroviral therapy might lower the performance of CD4 cell counts in classifying subject's t-week survival. One possible explanation for this conclusion is that the prior therapy makes patients more homogeneous in survival time and CD4 cell counts. For long term survival classification, the performance of CD4 cell counts is irrelevant to whether patients receive prior therapy or not. Since no significant difference between $\theta^{(1)}_t(q^{(1)}_{\alpha t})$ and $\theta^{(2)}_t(q^{(2)}_{\alpha t})$ exists, it would necessitate increasing sample size to detect a difference between the time-dependent PAUCs.

To characterize the effect of the therapy status ($Z = 1$ indicates a patient with prior therapy and 0 otherwise) on the classification accuracy of CD4 cell counts, we consider a saturated time-varying coefficient logistic regression model

$$\theta_t(Z_i, Z_j) = \frac{\exp(\gamma_{t1} + \gamma_{t2}Z_i + \gamma_{t3}Z_j + \gamma_{t4}Z_iZ_j)}{1 + \exp(\gamma_{t1} + \gamma_{t2}Z_i + \gamma_{t3}Z_j + \gamma_{t4}Z_iZ_j)}.i \neq j,$$

As evidenced by our numerical study, the performance of this working model is generally good in the estimation of $\theta_t(Z_i, Z_j)$. A positive linear combination of coefficients

$\gamma_{t2} + \gamma_{t3} + \gamma_{t4}$ (figure 5.3 (b)) means that the CD4 cell counts of patients with a prior therapy are more sensitive in classifying failure status than those without. We find that CD4 cell counts perform better among patients with a prior therapy except for weeks 118-152. However, the difference is not significant enough and the same conclusion can be drawn based on the nonparametric estimate of $\widetilde{\gamma}_\alpha(t)$. Using the time-varying coefficient logistic regression model, the estimated time-dependent AUCs are given in figures 5.3 (c)-(d). These figures show the appropriateness of model specification because the same patter as obtained by nonparametric methods (figures 5.1 (e)-(f)).

5.2.2 Angiography CAD Study

The main objective of this study is to evaluate the classification abilities of CRP, SAA, IL-6, and tHcy on the CAD-related vital status over time. Let $\theta_t^{(\text{CRP})}$, $\theta_t^{(\text{SAA})}$, $\theta_t^{(\text{IL-6})}$, and $\theta_t^{(\text{tHcy})}$ denote the time-dependent AUCs of CRP, SAA, IL-6, and tHcy, respectively. Nonparametric estimates of the quantities $(\widetilde{\theta}_t^{(\text{CRP})}, \widetilde{\theta}_t^{(\text{SAA})}, \widetilde{\theta}_t^{(\text{IL-6})}, \widetilde{\theta}_t^{(\text{tHcy})})$ based on $\widetilde{\theta}_t(\widetilde{q}_{1t})$ and $(\widehat{\theta}_t^{(\text{CRP})}, \widehat{\theta}_t^{(\text{SAA})}, \widehat{\theta}_t^{(\text{IL-6})}, \widehat{\theta}_t^{(\text{tHcy})})$ based on $\widehat{\theta}_t(\widehat{q}_{1t})$, which are computed under the validity of totally independent censoring and marker-dependent censoring, are displayed in figures 5.4-5.5. The estimates of $\theta_t^{(\text{tHcy})}$ is the highest at the beginning and decreases rapidly from 0.732 to 0.643, and the estimates of other biomarkers are all lower than 0.7. The comparable estimates are found in these figures and imply the first estimation approach is robust against violation of marker-dependent censoring. The results of statistical analysis in the paper of Lee, et al. (2006) indicated that the performance of IL-6 is superior than the other biomarkers. This conclusion is also evidenced in our analysis in which the classification accuracy of IL-6 is the largestat the end of study. We further detect that the plasma biomarker of tHcy has the best classification power at the stating time period. It can be seen that the time-dependent AUC of CRP and tHcy have a decreasing trend while that of IL-6 has an increasing trend. As for that of SAA, the estimates seem to stably stay around 0.6. The constructed 0.95 confidence intervals show that all plasma biomarkers are

indicative of disease over time.

To further improve the classification accuracy of multiple plasma biomarkers, we search for combinations of these biomarkers through a very flexible time-varying coefficient EGLM in (4.1) for the conditional survival distribution. In the numerical implementation, the linear predictors without considering tHcy are not likely to be optimal composite biomarker. For the sake of identifiability, the coefficient of tHcy in the linear predictor is set to be one. The smoothing parameters for $\widehat{\beta}_t$ are all set to be 0.05 and a bootstrap sampling is carried out to compute the standard errors of the estimators and construct the confidence intervals of the parameter functions. The analysis results are exhibited in tables 5.17-5.18. One can see that form table 5.17 that CRP and IL-6 are negatively and positively, respectively, associated with CAD-related death and significantly classify the CAD-related vital status after about day 1500 and day 2000. Except at the small period around day 2000, the effect of SAA is generally nonsignificant. Moreover, tHcy is detected to have an overall superior classification capacity within the study period. It is found in table 5.18 that the estimated values $\widetilde{\theta}_t(\widetilde{\beta}_t)$ and $\widehat{\theta}_t(\widehat{\beta}_t)$ decline as time progresses and are higher than 0.7 before day 3500, which demonstrate the advantage from combining biomarkers.

The appropriateness of a time-varying coefficient logistic regression model is also investigated via comparing the estimated linear predictors $\bar{\beta}_{t1}$tHcy+$\bar{\beta}_{t2}$CRP+$\bar{\beta}_{t3}$SAA+$\bar{\beta}_{t4}$IL-6 of Chiang and Huang (2009) with those of optimal linear predictors and the corresponding estimates of the time-dependent AUCs. The estimated coefficients imply a rather similar biological explanation. Although the estimated values $\widehat{\theta}_t(\bar{\beta}_t)$ and $\widetilde{\theta}_t(\bar{\beta}_t)$ are found to be relatively lower than $\widehat{\theta}_t(\widehat{\theta}_t)$ and $\widetilde{\theta}_t(\widetilde{\theta}_t)$, they are not significantly different from our estimates within the study period. The time-varying coefficient logistic regression model might be a suitable working model to characterize the conditional survival distribution.

Table 5.1: The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under totally independent censoring

$\alpha = 0.1 \quad c.r. = 30\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.0264	0.0276	0.0070	0.0048	0.826	0.0270	0.0049	0.0039	0.876
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0284	0.0071	0.0048	0.810	0.0275	0.0051	0.0038	0.846
(0.01)	$t_{0.6}$	0.0279	0.0293	0.0071	0.0050	0.808	0.0284	0.0053	0.0040	0.864
	$t_{0.4}$	0.0264	0.0221	0.0064	0.0078	0.942	0.0234	0.0047	0.0056	0.952
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0233	0.0068	0.0074	0.918	0.0244	0.0048	0.0054	0.942
(λ_{opt})	$t_{0.6}$	0.0279	0.0252	0.0076	0.0072	0.886	0.0258	0.0052	0.0054	0.910
	$t_{0.4}$	0.0264	0.0193	0.0043	0.0082	0.966	0.0189	0.0029	0.0059	0.872
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0210	0.0051	0.0077	0.932	0.0202	0.0036	0.0057	0.846
(0.2)	$t_{0.6}$	0.0279	0.0228	0.0059	0.0076	0.914	0.0218	0.0041	0.0058	0.854
	$t_{0.4}$	0.0264	0.0293	0.0070	0.0069	0.928	0.0279	0.0048	0.0048	0.928
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0306	0.0076	0.0071	0.880	0.0287	0.0049	0.0050	0.918
	$t_{0.6}$	0.0279	0.0332	0.0088	0.0075	0.832	0.0301	0.0056	0.0054	0.910

$\alpha = 0.1 \quad c.r. = 50\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.0264	0.0272	0.0073	0.0049	0.802	0.0268	0.0051	0.0040	0.880
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0271	0.0075	0.0050	0.786	0.0271	0.0056	0.0040	0.820
(0.01)	$t_{0.6}$	0.0279	0.0268	0.0082	0.0051	0.756	0.0278	0.0062	0.0042	0.798
	$t_{0.4}$	0.0264	0.0223	0.0070	0.0085	0.944	0.0234	0.0055	0.0062	0.944
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0235	0.0077	0.0081	0.910	0.0243	0.0057	0.0060	0.920
(λ_{opt})	$t_{0.6}$	0.0279	0.0251	0.0086	0.0080	0.884	0.0257	0.0060	0.0059	0.906
	$t_{0.4}$	0.0264	0.0192	0.0046	0.0091	0.964	0.0186	0.0032	0.0066	0.894
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0209	0.0055	0.0087	0.958	0.0200	0.0040	0.0064	0.878
(0.2)	$t_{0.6}$	0.0279	0.0227	0.0069	0.0085	0.914	0.0221	0.0049	0.0065	0.896
	$t_{0.4}$	0.0264	0.0304	0.0080	0.0081	0.922	0.0284	0.0060	0.0055	0.916
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0319	0.0091	0.0083	0.882	0.0292	0.0062	0.0057	0.920
	$t_{0.6}$	0.0279	0.0349	0.0106	0.0091	0.830	0.0312	0.0063	0.0062	0.908

Table 5.2: The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under totally independent censoring

$\alpha = 0.3 \quad c.r. = 30\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.1420	0.1436	0.0192	0.0138	0.836	0.1433	0.0134	0.0115	0.908
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1441	0.0190	0.0135	0.842	0.1434	0.0138	0.0111	0.886
(0.01)	$t_{0.6}$	0.1446	0.1451	0.0195	0.0138	0.820	0.1450	0.0140	0.0112	0.880
	$t_{0.4}$	0.1420	0.1331	0.0197	0.0215	0.958	0.1366	0.0136	0.0149	0.960
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1341	0.0202	0.0203	0.940	0.1376	0.0136	0.0142	0.960
(λ_{opt})	$t_{0.6}$	0.1446	0.1376	0.0210	0.0201	0.924	0.1403	0.0144	0.0143	0.940
	$t_{0.4}$	0.1420	0.1266	0.0164	0.0230	0.974	0.1257	0.0115	0.0166	0.920
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1294	0.0172	0.0215	0.954	0.1277	0.0125	0.0157	0.902
(0.2)	$t_{0.6}$	0.1446	0.1329	0.0183	0.0211	0.956	0.1309	0.0131	0.0157	0.914
	$t_{0.4}$	0.1420	0.1474	0.0187	0.0192	0.936	0.1448	0.0130	0.0134	0.952
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1480	0.0194	0.0191	0.928	0.1456	0.0135	0.0135	0.934
	$t_{0.6}$	0.1446	0.1523	0.0210	0.0199	0.900	0.1483	0.0145	0.0141	0.928

$\alpha = 0.3 \quad c.r. = 50\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.1420	0.1422	0.0211	0.0145	0.820	0.1423	0.0140	0.0120	0.908
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1414	0.0207	0.0142	0.814	0.1420	0.0148	0.0117	0.874
(0.01)	$t_{0.6}$	0.1446	0.1396	0.0233	0.0146	0.762	0.1437	0.0162	0.0119	0.848
	$t_{0.4}$	0.1420	0.1325	0.0217	0.0238	0.956	0.1355	0.0156	0.0166	0.952
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1345	0.0228	0.0228	0.932	0.1366	0.0157	0.0161	0.946
(λ_{opt})	$t_{0.6}$	0.1446	0.1372	0.0235	0.0226	0.928	0.1391	0.0161	0.0162	0.948
	$t_{0.4}$	0.1420	0.1263	0.0180	0.0257	0.972	0.1246	0.0125	0.0186	0.944
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1296	0.0186	0.0244	0.968	0.1269	0.0135	0.0178	0.930
(0.2)	$t_{0.6}$	0.1446	0.1325	0.0217	0.0241	0.954	0.1317	0.0152	0.0179	0.926
	$t_{0.4}$	0.1420	0.1477	0.0217	0.0220	0.942	0.1448	0.0156	0.0154	0.938
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1503	0.0229	0.0226	0.926	0.1458	0.0160	0.0158	0.938
	$t_{0.6}$	0.1446	0.1545	0.0247	0.0238	0.904	0.1490	0.0163	0.0168	0.940

Table 5.3: The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under totally independent censoring

$\alpha = 1 \quad c.r. = 30\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.7769	0.7765	0.0345	0.0255	0.842	0.7773	0.0236	0.0211	0.912
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7742	0.0341	0.0250	0.832	0.7748	0.0242	0.0205	0.884
(0.01)	$t_{0.6}$	0.7765	0.7732	0.0345	0.0253	0.854	0.7758	0.0242	0.0205	0.892
	$t_{0.4}$	0.7769	0.7574	0.0374	0.0388	0.962	0.7660	0.0259	0.0263	0.946
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7541	0.0386	0.0381	0.946	0.7639	0.0251	0.0258	0.954
(λ_{opt})	$t_{0.6}$	0.7765	0.7565	0.0385	0.0389	0.940	0.7648	0.0265	0.0265	0.948
	$t_{0.4}$	0.7769	0.7455	0.0334	0.0419	0.964	0.7451	0.0232	0.0299	0.914
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7447	0.0332	0.0409	0.968	0.7431	0.0241	0.0295	0.904
(0.2)	$t_{0.6}$	0.7765	0.7456	0.0341	0.0420	0.964	0.7445	0.0242	0.0303	0.902
	$t_{0.4}$	0.7769	0.7800	0.0326	0.0340	0.946	0.7790	0.0237	0.0239	0.944
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7766	0.0343	0.0339	0.946	0.7770	0.0237	0.0238	0.948
	$t_{0.6}$	0.7765	0.7794	0.0345	0.0349	0.934	0.7781	0.0250	0.0246	0.938

$\alpha = 1 \quad c.r. = 50\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.7769	0.7745	0.0378	0.0266	0.838	0.7765	0.0258	0.0221	0.896
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7712	0.0371	0.0261	0.818	0.7733	0.0267	0.0215	0.890
(0.01)	$t_{0.6}$	0.7765	0.7651	0.0414	0.0268	0.782	0.7738	0.0282	0.0216	0.880
	$t_{0.4}$	0.7769	0.7566	0.0413	0.0434	0.958	0.7639	0.0278	0.0297	0.962
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7553	0.0426	0.0428	0.948	0.7614	0.0290	0.0295	0.954
(λ_{opt})	$t_{0.6}$	0.7765	0.7538	0.0437	0.0444	0.956	0.7619	0.0299	0.0306	0.952
	$t_{0.4}$	0.7769	0.7451	0.0369	0.0467	0.968	0.7437	0.0255	0.0337	0.936
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7460	0.0362	0.0460	0.974	0.7423	0.0262	0.0334	0.932
(0.2)	$t_{0.6}$	0.7765	0.7457	0.0403	0.0478	0.976	0.7459	0.0277	0.0347	0.944
	$t_{0.4}$	0.7769	0.7785	0.0379	0.0390	0.946	0.7775	0.0267	0.0274	0.962
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7772	0.0393	0.0395	0.946	0.7749	0.0282	0.0279	0.936
	$t_{0.6}$	0.7765	0.7762	0.0426	0.0416	0.942	0.7755	0.0291	0.0293	0.946

Table 5.4: The empirical coverage probabilities of 0.95 simultaneous confidence bands under totally independent censoring

$c.r. = 30\%$		$n = 250$		$n = 500$	
	α	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$
	0.1	0.718	0.664	0.812	0.792
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.794	0.750	0.862	0.840
(0.01)	1	0.818	0.792	0.904	0.878
	0.1	0.916	0.882	0.932	0.930
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.946	0.930	0.962	0.952
(λ_{opt})	1	0.956	0.942	0.952	0.958
	0.1	0.930	0.920	0.882	0.866
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.962	0.960	0.932	0.928
(0.2)	1	0.966	0.964	0.918	0.924
	0.1	0.860	0.812	0.906	0.884
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	0.3	0.922	0.892	0.928	0.918
	1	0.940	0.924	0.926	0.924

$c.r. = 50\%$		$n = 250$		$n = 500$	
	α	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$
	0.1	0.714	0.650	0.772	0.712
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.728	0.672	0.832	0.806
(0.01)	1	0.760	0.722	0.854	0.824
	0.1	0.908	0.876	0.914	0.892
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.936	0.916	0.936	0.934
(λ_{opt})	1	0.960	0.952	0.962	0.956
	0.1	0.950	0.922	0.884	0.896
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.974	0.970	0.928	0.926
(0.2)	1	0.978	0.980	0.948	0.954
	0.1	0.850	0.788	0.868	0.854
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	0.3	0.898	0.884	0.928	0.918
	1	0.946	0.932	0.940	0.942

Table 5.5: The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under marker-dependent censoring

$\alpha = 0.1 \quad c.r. = 30\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.0264	0.0274	0.0073	0.0048	0.796	0.0268	0.0052	0.0040	0.834
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0276	0.0073	0.0048	0.794	0.0272	0.0055	0.0040	0.846
(0.01)	$t_{0.6}$	0.0279	0.0282	0.0080	0.0049	0.738	0.0279	0.0055	0.0041	0.848
	$t_{0.4}$	0.0264	0.0219	0.0068	0.0081	0.954	0.0234	0.0050	0.0058	0.948
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0233	0.0070	0.0077	0.926	0.0248	0.0054	0.0056	0.932
(λ_{opt})	$t_{0.6}$	0.0279	0.0252	0.0078	0.0075	0.890	0.0264	0.0055	0.0056	0.924
	$t_{0.4}$	0.0264	0.0192	0.0043	0.0086	0.974	0.0190	0.0032	0.0062	0.894
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0208	0.0053	0.0082	0.942	0.0203	0.0037	0.0061	0.868
(0.2)	$t_{0.6}$	0.0279	0.0231	0.0067	0.0082	0.918	0.0218	0.0045	0.0061	0.864
	$t_{0.4}$	0.0264	0.0301	0.0076	0.0082	0.940	0.0294	0.0052	0.0058	0.948
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0314	0.0077	0.0085	0.928	0.0301	0.0053	0.0060	0.932
	$t_{0.6}$	0.0279	0.0340	0.0084	0.0092	0.902	0.0314	0.0058	0.0065	0.940

$\alpha = 0.1 \quad c.r. = 50\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.0264	0.0258	0.0080	0.0050	0.760	0.0266	0.0062	0.0042	0.786
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0248	0.0079	0.0049	0.722	0.0268	0.0064	0.0041	0.772
(0.01)	$t_{0.6}$	0.0279	0.0238	0.0083	0.0050	0.668	0.0265	0.0066	0.0043	0.754
	$t_{0.4}$	0.0264	0.0233	0.0078	0.0093	0.954	0.0237	0.0057	0.0067	0.950
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0245	0.0086	0.0086	0.912	0.0250	0.0061	0.0064	0.918
(λ_{opt})	$t_{0.6}$	0.0279	0.0268	0.0095	0.0081	0.864	0.0266	0.0071	0.0064	0.892
	$t_{0.4}$	0.0264	0.0195	0.0053	0.0098	0.970	0.0192	0.0037	0.0073	0.952
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0214	0.0066	0.0094	0.928	0.0209	0.0047	0.0071	0.932
(0.2)	$t_{0.6}$	0.0279	0.0235	0.0078	0.0090	0.906	0.0225	0.0056	0.0072	0.908
	$t_{0.4}$	0.0264	0.0337	0.0090	0.0116	0.934	0.0303	0.0060	0.0080	0.972
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.0269	0.0346	0.0093	0.0120	0.920	0.0310	0.0062	0.0085	0.986
	$t_{0.6}$	0.0279	0.0380	0.0099	0.0134	0.928	0.0330	0.0067	0.0095	0.978

Table 5.6: The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under marker-dependent censoring

$\alpha = 0.3 \quad c.r. = 30\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.1420	0.1431	0.0197	0.0141	0.836	0.1424	0.0140	0.0116	0.880
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1424	0.0197	0.0138	0.844	0.1427	0.0142	0.0113	0.874
(0.01)	$t_{0.6}$	0.1446	0.1429	0.0204	0.0141	0.830	0.1438	0.0137	0.0115	0.892
	$t_{0.4}$	0.1420	0.1318	0.0210	0.0223	0.952	0.1362	0.0144	0.0154	0.948
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1337	0.0208	0.0210	0.938	0.1383	0.0150	0.0147	0.932
(λ_{opt})	$t_{0.6}$	0.1446	0.1374	0.0213	0.0207	0.930	0.1414	0.0143	0.0146	0.940
	$t_{0.4}$	0.1420	0.1272	0.0169	0.0242	0.982	0.1265	0.0122	0.0174	0.938
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1297	0.0179	0.0227	0.960	0.1281	0.0129	0.0164	0.914
(0.2)	$t_{0.6}$	0.1446	0.1342	0.0191	0.0223	0.956	0.1310	0.0137	0.0164	0.916
	$t_{0.4}$	0.1420	0.1493	0.0200	0.0228	0.946	0.1496	0.0139	0.0159	0.950
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1504	0.0194	0.0221	0.946	0.1494	0.0138	0.0156	0.952
	$t_{0.6}$	0.1446	0.1543	0.0197	0.0225	0.950	0.1515	0.0145	0.0159	0.930

$\alpha = 0.3 \quad c.r. = 50\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.1420	0.1377	0.0226	0.0148	0.788	0.1417	0.0167	0.0122	0.826
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1330	0.0222	0.0145	0.768	0.1407	0.0168	0.0118	0.824
(0.01)	$t_{0.6}$	0.1446	0.1274	0.0231	0.0147	0.668	0.1389	0.0176	0.0120	0.796
	$t_{0.4}$	0.1420	0.1355	0.0238	0.0249	0.942	0.1369	0.0168	0.0176	0.948
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1366	0.0242	0.0235	0.928	0.1390	0.0173	0.0168	0.934
(λ_{opt})	$t_{0.6}$	0.1446	0.1410	0.0247	0.0228	0.916	0.1418	0.0184	0.0169	0.928
	$t_{0.4}$	0.1420	0.1278	0.0196	0.0275	0.974	0.1273	0.0137	0.0201	0.968
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1306	0.0212	0.0257	0.960	0.1295	0.0154	0.0190	0.954
(0.2)	$t_{0.6}$	0.1446	0.1340	0.0227	0.0249	0.948	0.1321	0.0163	0.0190	0.938
	$t_{0.4}$	0.1420	0.1569	0.0232	0.0297	0.948	0.1523	0.0157	0.0212	0.964
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.1423	0.1570	0.0231	0.0286	0.946	0.1521	0.0154	0.0203	0.958
	$t_{0.6}$	0.1446	0.1615	0.0234	0.0284	0.936	0.1541	0.0163	0.0203	0.944

Table 5.7: The averages (Mean) and the standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) under marker-dependent censoring

$\alpha = 1 \quad c.r. = 30\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.7769	0.7775	0.0355	0.0256	0.874	0.7760	0.0244	0.0212	0.926
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7729	0.0352	0.0252	0.844	0.7746	0.0238	0.0205	0.906
(0.01)	$t_{0.6}$	0.7765	0.7708	0.0353	0.0256	0.858	0.7743	0.0247	0.0207	0.886
	$t_{0.4}$	0.7769	0.7555	0.0398	0.0396	0.940	0.7665	0.0263	0.0268	0.948
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7531	0.0391	0.0387	0.934	0.7631	0.0263	0.0264	0.946
(λ_{opt})	$t_{0.6}$	0.7765	0.7557	0.0385	0.0392	0.946	0.7633	0.0269	0.0270	0.944
	$t_{0.4}$	0.7769	0.7479	0.0341	0.0426	0.968	0.7436	0.0237	0.0305	0.888
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7461	0.0341	0.0415	0.956	0.7431	0.0236	0.0298	0.894
(0.2)	$t_{0.6}$	0.7765	0.7486	0.0339	0.0422	0.968	0.7442	0.0242	0.0306	0.908
	$t_{0.4}$	0.7769	0.7827	0.0353	0.0345	0.918	0.7853	0.0242	0.0241	0.920
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7792	0.0342	0.0342	0.942	0.7812	0.0240	0.0239	0.930
	$t_{0.6}$	0.7765	0.7811	0.0336	0.0349	0.950	0.7809	0.0246	0.0245	0.918

$\alpha = 1 \quad c.r. = 50\%$

n			250				500			
	t_q	$\theta_t(q_{\alpha t})$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.4}$	0.7769	0.7658	0.0419	0.0272	0.788	0.7755	0.0280	0.0222	0.866
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7554	0.0408	0.0271	0.760	0.7682	0.0288	0.0218	0.854
(0.01)	$t_{0.6}$	0.7765	0.7418	0.0435	0.0281	0.660	0.7621	0.0308	0.0223	0.824
	$t_{0.4}$	0.7769	0.7605	0.0438	0.0434	0.944	0.7650	0.0298	0.0304	0.956
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7576	0.0441	0.0425	0.932	0.7621	0.0303	0.0301	0.942
(λ_{opt})	$t_{0.6}$	0.7765	0.7607	0.0449	0.0429	0.944	0.7626	0.0323	0.0310	0.934
	$t_{0.4}$	0.7769	0.7465	0.0393	0.0481	0.972	0.7457	0.0277	0.0346	0.930
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7455	0.0399	0.0467	0.968	0.7424	0.0280	0.0342	0.928
(0.2)	$t_{0.6}$	0.7765	0.7454	0.0413	0.0475	0.964	0.7420	0.0296	0.0351	0.924
	$t_{0.4}$	0.7769	0.7915	0.0398	0.0388	0.914	0.7879	0.0273	0.0276	0.936
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	$t_{0.5}$	0.7746	0.7859	0.0407	0.0388	0.908	0.7823	0.0268	0.0276	0.938
	$t_{0.6}$	0.7765	0.7871	0.0413	0.0397	0.928	0.7808	0.0285	0.0284	0.942

Table 5.8: The empirical coverage probabilities of 0.95 simultaneous confidence bands under marker-dependent censoring

$c.r. = 30\%$		$n = 250$		$n = 500$	
	α	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$
	0.1	0.710	0.606	0.800	0.754
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.780	0.734	0.860	0.838
(0.01)	1	0.834	0.802	0.890	0.868
	0.1	0.918	0.878	0.936	0.908
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.932	0.918	0.946	0.936
(λ_{opt})	1	0.946	0.942	0.948	0.946
	0.1	0.946	0.924	0.888	0.878
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.970	0.966	0.934	0.936
(0.2)	1	0.970	0.976	0.914	0.922
	0.1	0.912	0.888	0.912	0.900
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	0.3	0.938	0.932	0.944	0.940
	1	0.914	0.922	0.910	0.908

$c.r. = 50\%$		$n = 250$		$n = 500$	
	α	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$	$[t_{0.4}, t_{0.5}]$	$[t_{0.4}, t_{0.6}]$
	0.1	0.616	0.520	0.702	0.614
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.670	0.584	0.774	0.720
(0.01)	1	0.712	0.610	0.850	0.794
	0.1	0.916	0.844	0.908	0.878
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.926	0.900	0.940	0.928
(λ_{opt})	1	0.940	0.924	0.958	0.954
	0.1	0.938	0.896	0.940	0.896
$\widehat{\theta}_t(\widehat{q}_{\alpha t})$	0.3	0.964	0.948	0.966	0.952
(0.2)	1	0.972	0.968	0.946	0.942
	0.1	0.920	0.904	0.970	0.972
$\widetilde{\theta}_t(\widetilde{q}_{\alpha t})$	0.3	0.950	0.932	0.968	0.952
	1	0.884	0.878	0.918	0.916

Table 5.9: The averages (Mean) and standard deviations (SD) of 500 estimates, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) of 0.95 pointwise confidence intervals of $\widehat{\gamma}_t$

$c.r. = 30\%$

n			250				500			
	t_q	γ_t	Mean	SD	SE	CP	Mean	SD	SE	CP
$\widehat{\gamma}_{t1}$	$t_{0.3}$	0.086	0.088	0.054	0.053	0.948	0.086	0.037	0.037	0.942
	$t_{0.5}$	0.111	0.112	0.051	0.054	0.952	0.112	0.038	0.038	0.938
	$t_{0.7}$	0.157	0.162	0.067	0.070	0.944	0.161	0.047	0.049	0.940
$\widehat{\gamma}_{t2}$	$t_{0.3}$	0.115	0.117	0.083	0.080	0.928	0.114	0.054	0.055	0.948
	$t_{0.5}$	0.099	0.101	0.079	0.080	0.956	0.098	0.054	0.054	0.960
	$t_{0.7}$	0.063	0.058	0.103	0.105	0.946	0.063	0.071	0.072	0.954
$\widehat{\gamma}_{t3}$	$t_{0.3}$	0.427	0.425	0.104	0.101	0.932	0.430	0.073	0.071	0.948
	$t_{0.5}$	0.469	0.463	0.104	0.102	0.930	0.473	0.073	0.072	0.944
	$t_{0.7}$	0.518	0.521	0.123	0.119	0.912	0.521	0.084	0.083	0.940
$\widehat{\gamma}_{t4}$	$t_{0.3}$	0.286	0.284	0.111	0.108	0.940	0.284	0.076	0.075	0.940
	$t_{0.5}$	0.211	0.212	0.110	0.108	0.938	0.208	0.074	0.075	0.954
	$t_{0.7}$	0.105	0.099	0.131	0.125	0.934	0.100	0.093	0.088	0.908

$c.r. = 50\%$

			n=250				n=500			
	t_q	γ_t	Mean	SD	SE	CP	Mean	SD	SE	CP
$\widehat{\gamma}_{t1}$	$t_{0.3}$	0.086	0.089	0.057	0.057	0.932	0.085	0.038	0.040	0.946
	$t_{0.5}$	0.111	0.115	0.064	0.063	0.934	0.108	0.041	0.044	0.942
	$t_{0.7}$	0.157	0.151	0.106	0.099	0.904	0.156	0.073	0.069	0.912
$\widehat{\gamma}_{t2}$	$t_{0.3}$	0.115	0.109	0.085	0.085	0.928	0.116	0.057	0.059	0.954
	$t_{0.5}$	0.099	0.096	0.098	0.095	0.936	0.102	0.063	0.066	0.944
	$t_{0.7}$	0.063	0.075	0.177	0.155	0.910	0.066	0.111	0.108	0.934
$\widehat{\gamma}_{t3}$	$t_{0.3}$	0.427	0.418	0.109	0.110	0.952	0.429	0.073	0.076	0.954
	$t_{0.5}$	0.469	0.470	0.120	0.122	0.950	0.471	0.082	0.085	0.962
	$t_{0.7}$	0.518	0.519	0.179	0.166	0.910	0.517	0.119	0.120	0.924
$\widehat{\gamma}_{t4}$	$t_{0.3}$	0.286	0.301	0.122	0.117	0.922	0.287	0.081	0.082	0.952
	$t_{0.5}$	0.211	0.216	0.132	0.131	0.950	0.210	0.091	0.091	0.944
	$t_{0.7}$	0.105	0.112	0.196	0.179	0.906	0.108	0.133	0.131	0.930

Table 5.10: The averages (Mean) and standard deviations (SD) of 500 estimates $\widehat{\theta}_t^{(T)}(z_0, z_0)$'s, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) of 0.95 pointwise confidence intervals

$c.r. = 30\%$

n			250				500			
z	t_q	$\theta(t\|z,z)$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.3}$	0.560	0.560	0.038	0.038	0.934	0.561	0.026	0.027	0.950
1.5	$t_{0.5}$	0.571	0.570	0.038	0.038	0.950	0.573	0.027	0.027	0.938
	$t_{0.7}$	0.587	0.590	0.046	0.047	0.952	0.591	0.034	0.033	0.948
	$t_{0.3}$	0.613	0.612	0.052	0.050	0.932	0.615	0.035	0.035	0.948
2.5	$t_{0.5}$	0.634	0.632	0.051	0.050	0.940	0.637	0.036	0.035	0.944
	$t_{0.7}$	0.665	0.669	0.059	0.060	0.946	0.668	0.044	0.043	0.935

$c.r. = 50\%$

n			250				500			
z	t_q	$\theta(t\|z,z)$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.3}$	0.560	0.562	0.042	0.040	0.934	0.562	0.027	0.029	0.960
1.5	$t_{0.5}$	0.571	0.577	0.043	0.044	0.944	0.571	0.032	0.031	0.950
	$t_{0.7}$	0.587	0.590	0.071	0.066	0.888	0.588	0.051	0.048	0.910
	$t_{0.3}$	0.613	0.614	0.055	0.053	0.940	0.615	0.036	0.038	0.962
2.5	$t_{0.5}$	0.634	0.641	0.056	0.058	0.954	0.634	0.041	0.042	0.946
	$t_{0.7}$	0.665	0.665	0.093	0.086	0.894	0.665	0.064	0.062	0.932

Table 5.11: The averages (Mean) and standard deviations (SD) of 500 estimates $\widehat{\theta}_t^{(W)}(z_0, z_0)$'s, the averages of 500 standard errors (SE) of estimators, and the empirical coverage probabilities (CP) of 0.95 pointwise confidence intervals

$c.r. = 30\%$

n			250				500			
z	t_q	$\theta(t\|z,z)$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.3}$	0.560	0.567	0.047	0.047	0.926	0.567	0.033	0.033	0.944
1.5	$t_{0.5}$	0.571	0.576	0.045	0.045	0.950	0.579	0.033	0.032	0.925
	$t_{0.7}$	0.587	0.597	0.054	0.054	0.944	0.597	0.039	0.038	0.931
	$t_{0.3}$	0.613	0.624	0.059	0.057	0.906	0.626	0.040	0.040	0.931
2.5	$t_{0.5}$	0.634	0.641	0.055	0.054	0.936	0.646	0.039	0.038	0.913
	$t_{0.7}$	0.665	0.676	0.061	0.062	0.932	0.675	0.045	0.044	0.927

$c.r. = 50\%$

n			250				500			
z	t_q	$\theta(t\|z,z)$	Mean	SD	SE	CP	Mean	SD	SE	CP
	$t_{0.3}$	0.560	0.569	0.053	0.050	0.926	0.569	0.034	0.036	0.954
1.5	$t_{0.5}$	0.571	0.585	0.052	0.053	0.936	0.577	0.038	0.038	0.938
	$t_{0.7}$	0.587	0.597	0.083	0.077	0.888	0.595	0.059	0.055	0.916
	$t_{0.3}$	0.613	0.626	0.063	0.061	0.924	0.627	0.041	0.043	0.944
2.5	$t_{0.5}$	0.634	0.651	0.060	0.063	0.928	0.643	0.044	0.045	0.934
	$t_{0.7}$	0.665	0.673	0.096	0.088	0.882	0.672	0.066	0.063	0.920

Table 5.12: The empirical coverage probabilities of 0.95 simultaneous confidence bands with the sample sizes of 250 and 500, and the censoring rates of 30% and 50%

$c.r. = 30\%$

	$n = 250$		$n = 500$	
	$[t_{0.5}, t_{0.6}]$	$[t_{0.3}, t_{0.7}]$	$[t_{0.5}, t_{0.6}]$	$[t_{0.3}, t_{0.7}]$
$\widehat{\gamma}_{t1}$	0.936	0.926	0.946	0.933
$\widehat{\gamma}_{t2}$	0.964	0.934	0.958	0.950
$\widehat{\gamma}_{t3}$	0.926	0.918	0.923	0.931
$\widehat{\gamma}_{t4}$	0.928	0.910	0.946	0.921
$\widehat{\theta}_t^{(T)}(1.5, 1.5)$	0.940	0.922	0.933	0.927
$\widehat{\theta}_t^{(T)}(2.5, 2.5)$	0.932	0.920	0.935	0.948
$\widehat{\theta}_t^{(W)}(1.5, 1.5)$	0.942	0.922	0.921	0.938
$\widehat{\theta}_t^{(W)}(2.5, 2.5)$	0.910	0.886	0.896	0.915

$c.r. = 50\%$

	$n = 250$		$n = 500$	
	$[t_{0.5}, t_{0.6}]$	$[t_{0.3}, t_{0.7}]$	$[t_{0.5}, t_{0.6}]$	$[t_{0.3}, t_{0.7}]$
$\widehat{\gamma}_{t1}$	0.896	0.896	0.934	0.918
$\widehat{\gamma}_{t2}$	0.926	0.902	0.948	0.948
$\widehat{\gamma}_{t3}$	0.916	0.894	0.936	0.930
$\widehat{\gamma}_{t4}$	0.920	0.892	0.936	0.930
$\widehat{\theta}_t^{(T)}(1.5, 1.5)$	0.944	0.900	0.946	0.922
$\widehat{\theta}_t^{(T)}(2.5, 2.5)$	0.934	0.896	0.946	0.930
$\widehat{\theta}_t^{(W)}(1.5, 1.5)$	0.930	0.902	0.932	0.930
$\widehat{\theta}_t^{(W)}(2.5, 2.5)$	0.908	0.846	0.924	0.900

Table 5.13: The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under totally independent censoring with $c.r. = 30\%$

n	t_q	β_t	$\widehat{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widehat{\theta}_t(\widehat{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.693	0.050	0.050	0.95	0.817	0.816	0.034	0.033	0.95
250	$t_{0.5}$	0.637	0.648	0.038	0.042	0.96	0.803	0.801	0.035	0.033	0.96
	$t_{0.7}$	0.602	0.612	0.034	0.040	0.95	0.805	0.802	0.037	0.037	0.95
	$t_{0.3}$	0.679	0.689	0.035	0.034	0.93	0.817	0.812	0.024	0.024	0.95
500	$t_{0.5}$	0.637	0.646	0.027	0.028	0.92	0.803	0.798	0.024	0.024	0.96
	$t_{0.7}$	0.602	0.609	0.023	0.025	0.95	0.805	0.799	0.025	0.027	0.97

n	t_q	β_t	$\widetilde{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widetilde{\theta}_t(\widetilde{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.683	0.042	0.046	0.95	0.817	0.821	0.031	0.031	0.93
250	$t_{0.5}$	0.637	0.639	0.032	0.038	0.97	0.803	0.808	0.032	0.031	0.93
	$t_{0.7}$	0.602	0.604	0.029	0.034	0.97	0.805	0.812	0.035	0.035	0.92
	$t_{0.3}$	0.679	0.684	0.033	0.032	0.96	0.817	0.816	0.023	0.023	0.93
500	$t_{0.5}$	0.637	0.640	0.023	0.025	0.95	0.803	0.802	0.023	0.023	0.94
	$t_{0.7}$	0.602	0.603	0.021	0.022	0.93	0.805	0.807	0.024	0.025	0.94

Table 5.14: The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under totally independent censoring with $c.r. = 50\%$

n	t_q	β_t	$\widehat{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widehat{\theta}_t(\widehat{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.708	0.063	0.057	0.936	0.817	0.811	0.040	0.039	0.964
250	$t_{0.5}$	0.637	0.662	0.054	0.054	0.916	0.803	0.792	0.042	0.040	0.958
	$t_{0.7}$	0.602	0.627	0.052	0.054	0.922	0.805	0.792	0.046	0.045	0.962
	$t_{0.3}$	0.679	0.690	0.037	0.038	0.952	0.817	0.812	0.028	0.027	0.958
500	$t_{0.5}$	0.637	0.649	0.031	0.034	0.944	0.803	0.796	0.029	0.028	0.940
	$t_{0.7}$	0.602	0.614	0.028	0.033	0.928	0.805	0.794	0.032	0.032	0.932

n	t_q	β_t	$\widetilde{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widetilde{\theta}_t(\widetilde{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.686	0.051	0.053	0.966	0.817	0.823	0.037	0.036	0.924
250	$t_{0.5}$	0.637	0.643	0.040	0.046	0.962	0.803	0.808	0.037	0.037	0.934
	$t_{0.7}$	0.602	0.607	0.040	0.043	0.954	0.805	0.814	0.041	0.041	0.914
	$t_{0.3}$	0.679	0.681	0.035	0.036	0.970	0.817	0.820	0.026	0.026	0.932
500	$t_{0.5}$	0.637	0.638	0.026	0.030	0.970	0.803	0.807	0.028	0.026	0.922
	$t_{0.7}$	0.602	0.605	0.025	0.027	0.958	0.805	0.809	0.030	0.030	0.934

Table 5.15: The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under marker-dependent censoring with $c.r. = 30\%$

n	t_q	β_t	$\widehat{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widehat{\theta}_t(\widehat{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.692	0.054	0.054	0.95	0.817	0.811	0.037	0.037	0.96
250	$t_{0.5}$	0.637	0.645	0.047	0.045	0.96	0.803	0.796	0.037	0.036	0.96
	$t_{0.7}$	0.602	0.610	0.042	0.042	0.95	0.805	0.799	0.041	0.039	0.95
	$t_{0.3}$	0.679	0.687	0.037	0.037	0.97	0.817	0.813	0.027	0.026	0.96
500	$t_{0.5}$	0.637	0.643	0.030	0.030	0.94	0.803	0.799	0.028	0.026	0.95
	$t_{0.7}$	0.602	0.605	0.026	0.027	0.96	0.805	0.800	0.029	0.029	0.96

n	t_q	β_t	$\widetilde{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widetilde{\theta}_t(\widetilde{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.688	0.045	0.048	0.97	0.817	0.833	0.030	0.032	0.89
250	$t_{0.5}$	0.637	0.644	0.034	0.039	0.96	0.803	0.818	0.031	0.031	0.88
	$t_{0.7}$	0.602	0.609	0.032	0.035	0.96	0.805	0.819	0.034	0.033	0.90
	$t_{0.3}$	0.681	0.687	0.031	0.033	0.95	0.817	0.833	0.022	0.023	0.86
500	$t_{0.5}$	0.638	0.643	0.024	0.026	0.95	0.803	0.817	0.022	0.022	0.88
	$t_{0.7}$	0.605	0.607	0.020	0.023	0.96	0.805	0.819	0.024	0.024	0.87

Table 5.16: The averages and standard deviations (SD) of 500 estimates, the bootstrap standard errors (BSE), and the empirical coverage probabilities (CP) under marker-dependent censoring with $c.r. = 50\%$

n	t_q	β_t	$\widehat{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widehat{\theta}_t(\widehat{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.704	0.075	0.065	0.936	0.817	0.805	0.047	0.045	0.964
250	$t_{0.5}$	0.637	0.658	0.068	0.060	0.922	0.803	0.790	0.047	0.043	0.952
	$t_{0.7}$	0.602	0.619	0.057	0.057	0.928	0.806	0.790	0.049	0.047	0.958
	$t_{0.3}$	0.679	0.697	0.051	0.044	0.934	0.817	0.808	0.035	0.032	0.942
500	$t_{0.5}$	0.637	0.648	0.041	0.039	0.920	0.803	0.795	0.036	0.032	0.926
	$t_{0.7}$	0.602	0.611	0.037	0.036	0.938	0.805	0.797	0.039	0.034	0.930

n	t_q	β_t	$\widetilde{\beta}_t$	SD	BSE	CP	$\theta_t(\beta_t)$	$\widetilde{\theta}_t(\widetilde{\beta}_t)$	SD	BSE	CP
	$t_{0.3}$	0.679	0.691	0.057	0.055	0.964	0.817	0.842	0.036	0.036	0.822
250	$t_{0.5}$	0.637	0.645	0.044	0.046	0.948	0.803	0.827	0.036	0.035	0.816
	$t_{0.7}$	0.602	0.612	0.037	0.042	0.966	0.806	0.827	0.037	0.037	0.874
	$t_{0.3}$	0.681	0.690	0.038	0.038	0.962	0.817	0.841	0.026	0.026	0.804
500	$t_{0.5}$	0.638	0.645	0.028	0.031	0.968	0.803	0.824	0.026	0.025	0.798
	$t_{0.7}$	0.605	0.609	0.025	0.027	0.944	0.805	0.825	0.027	0.027	0.836

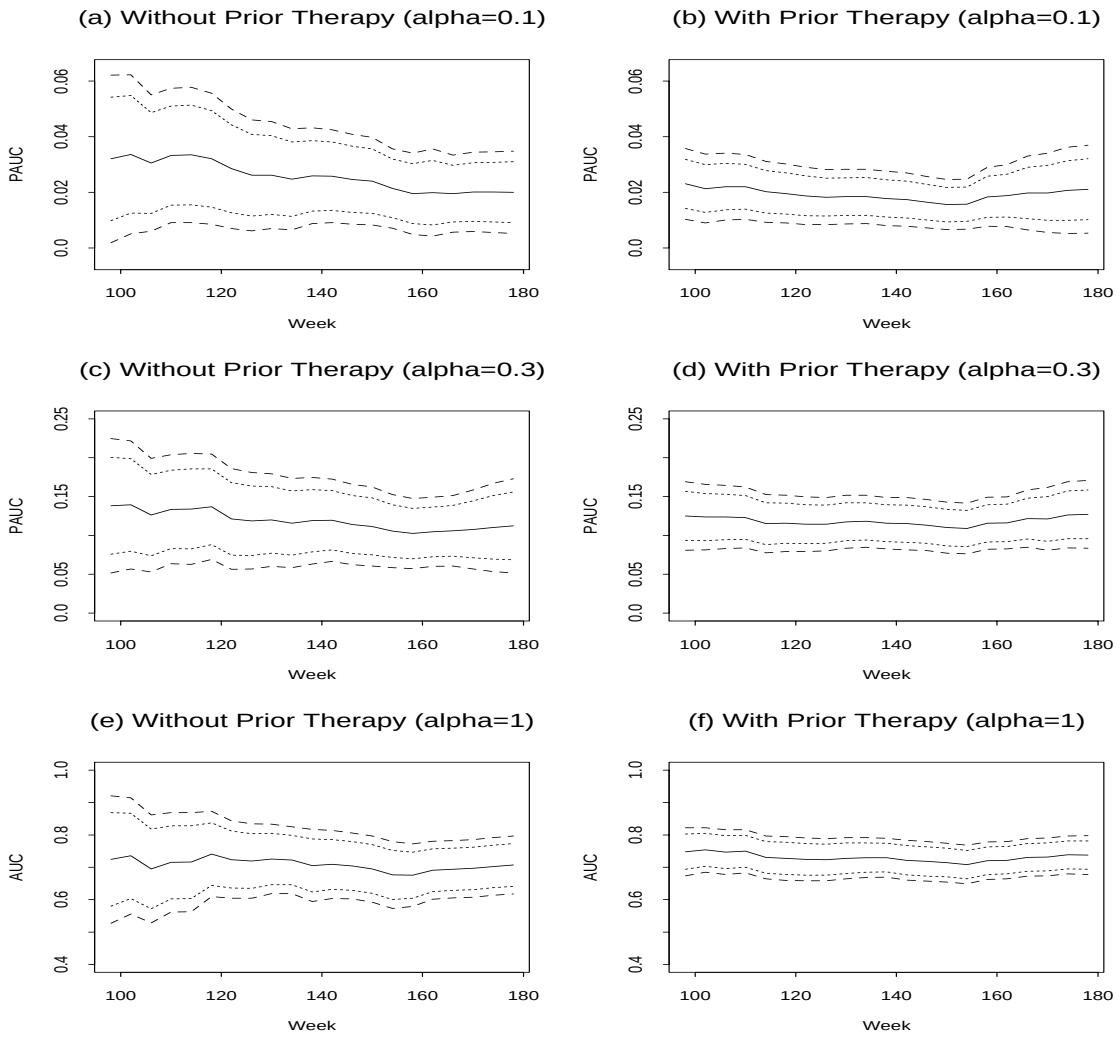

Figure 5.1: The estimated time-dependent PAUCs (solid curve) with the corresponding 0.95 pointwise (dotted curve) and simultaneous confidence bands (dashed curve)

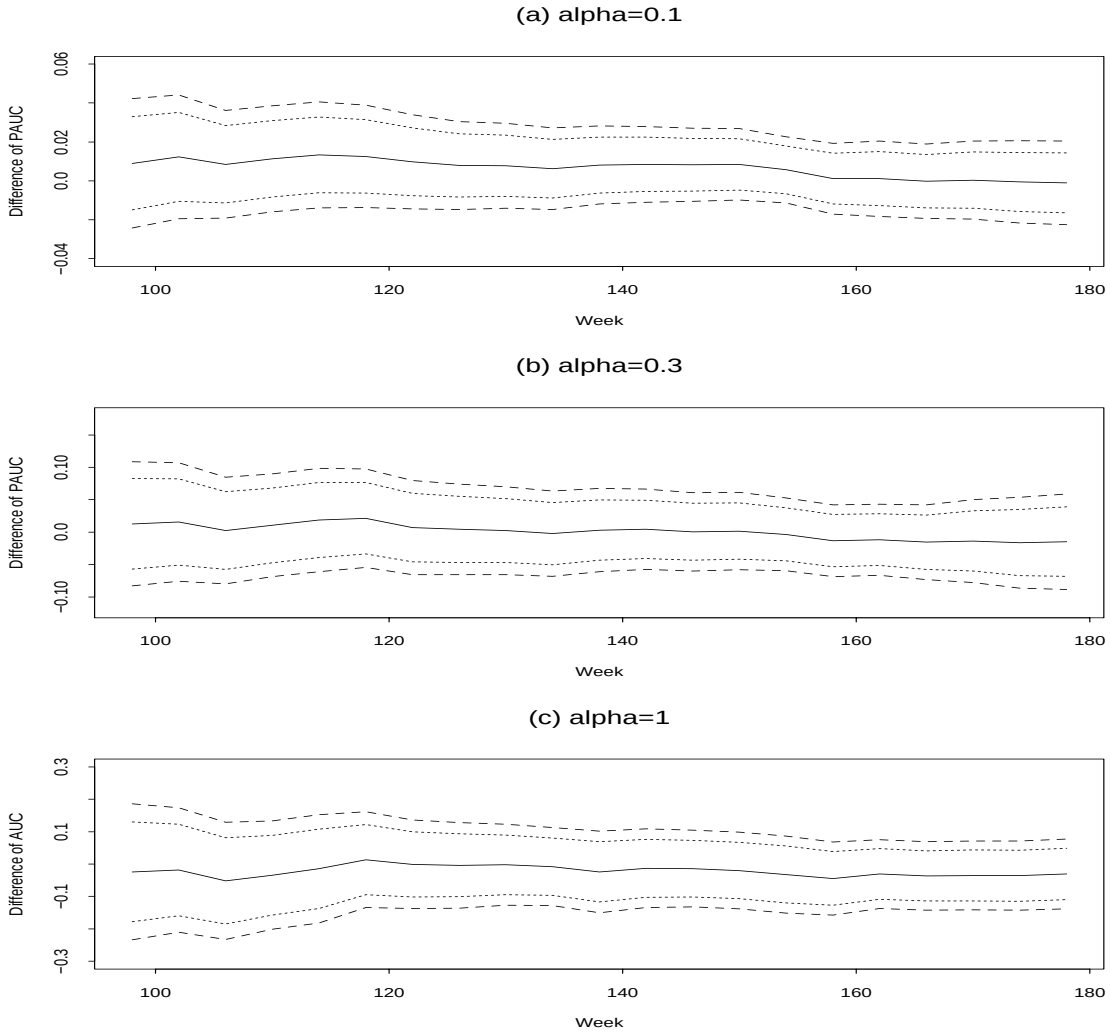

Figure 5.2: The estimated curves for the difference of the time-dependent PAUCs between non-therapy and therapy patients (solid curve) with the corresponding 0.95 pointwise (dotted curve) and simultaneous confidence bands (dashed curve)

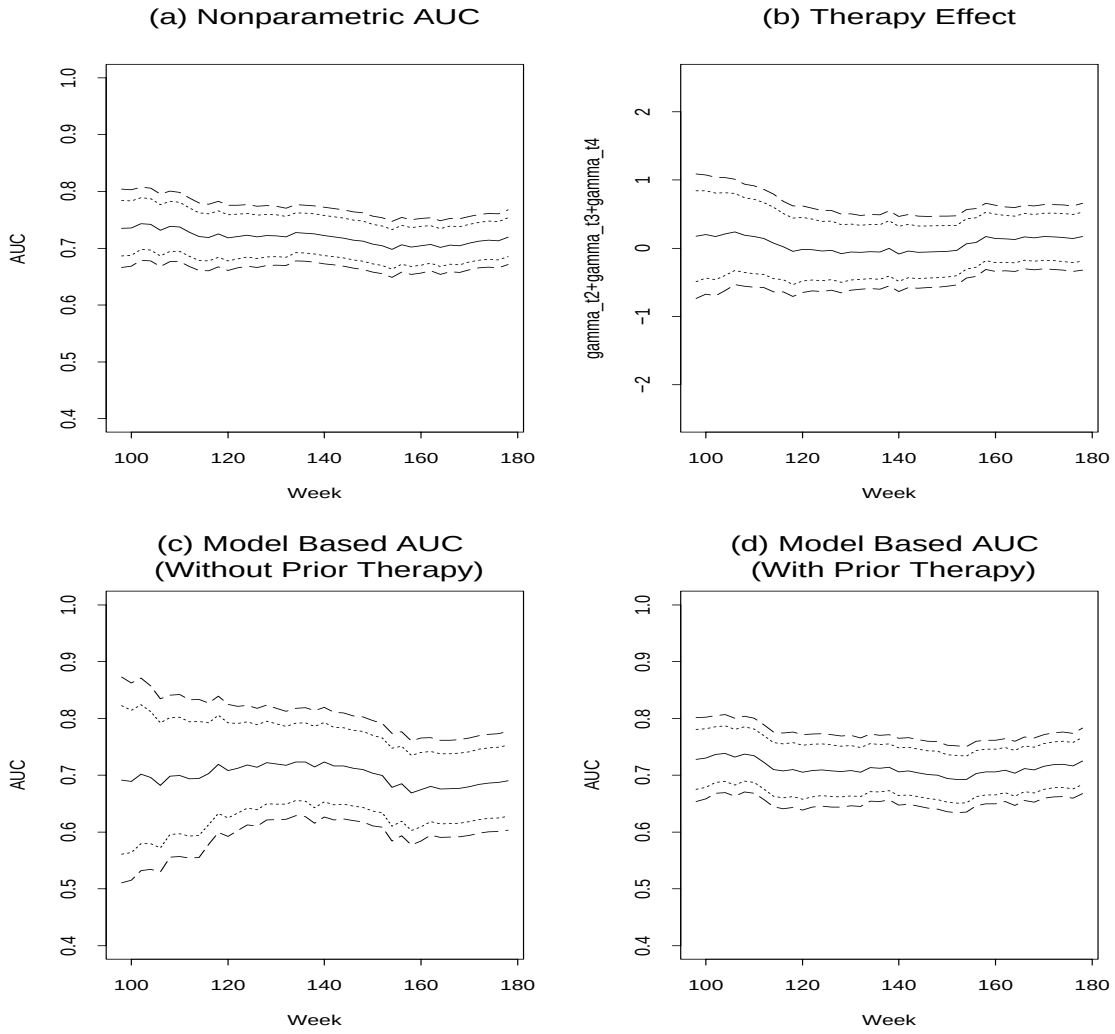

Figure 5.3: The estimated coefficient and time-dependent AUC curves (solid curve) with the corresponding 0.95 pointwise (dotted curve) and simultaneous confidence bands (dashed curve)

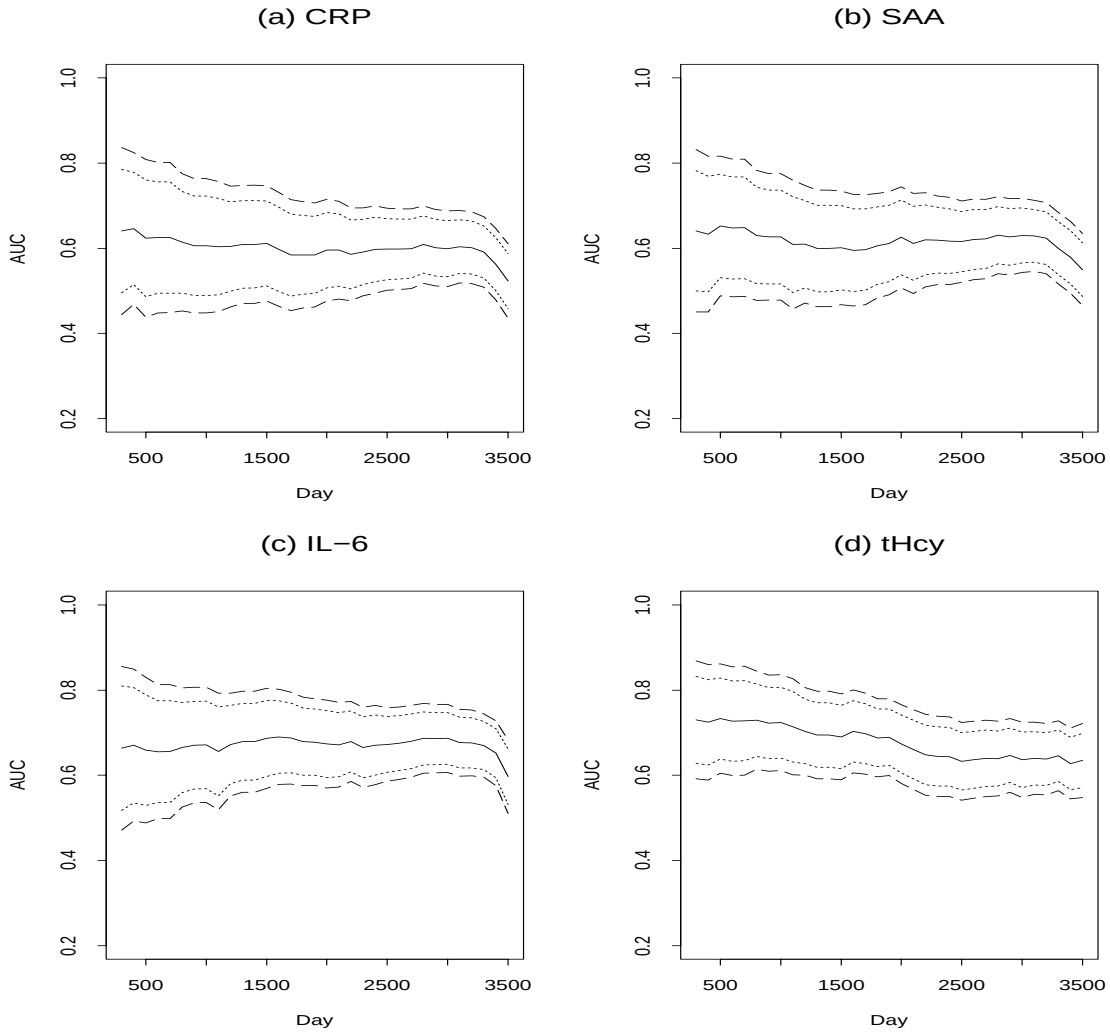

Figure 5.4: The estimated time-dependent AUCs (solid curve) under totally independent censoring with the corresponding 0.95 pointwise (dotted curve) and simultaneous (dashed curve) confidence bands

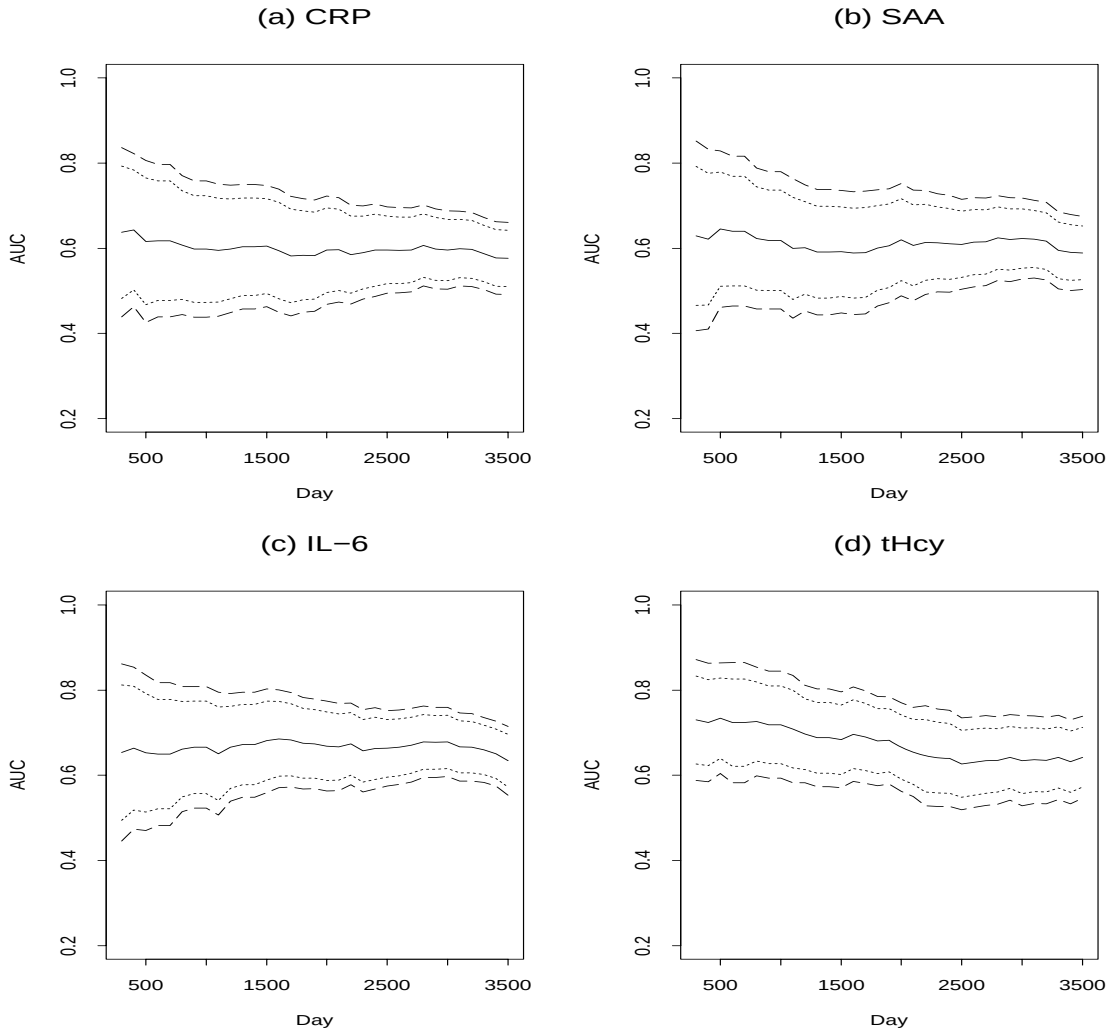

Figure 5.5: The estimated time-dependent AUCs (solid curve) under marker-dependent censoring with the corresponding 0.95 pointwise (dotted curve) and simultaneous (dashed curve) confidence bands

Table 5.17: The estimates of the coefficients in the linear predictor tHcy + β_{t2}CRP + β_{t3}SAA + β_{t4}IL-6 (95% bootstrap confidence intervals) at the selected time points

Days	$\widehat{\beta}_{t2}$	$\widehat{\beta}_{t3}$	$\widehat{\beta}_{t4}$
300	-0.25 (-1.64 , 1.69)	0.22 (-0.41 , 1.04)	-0.31 (-1.51 , 3.45)
500	-0.16 (-1.54 , 0.32)	0.27 (-0.45 , 1.60)	-0.33 (-1.23 , 3.13)
1000	-0.38 (-2.40 , 0.02)	0.23 (-0.39 , 1.63)	0.51 (-0.84 , 6.96)
1500	-1.32 (-4.72 , 0.04)	0.27 (-0.77 , 1.76)	4.76 (0.11 , 15.95)
2000	-1.79 (-5.67 , -0.43)	1.17 (0.01 , 3.13)	4.29 (-0.51 , 14.49)
2500	-2.32 (-6.62 , -0.60)	0.44 (-0.43 , 2.64)	9.74 (1.11 , 26.09)
3000	-4.31 (-8.67 , -1.00)	1.81 (-0.16 , 4.01)	14.43 (2.42 , 29.73)
3500	-1.70 (-6.02 , -0.50)	0.24 (-0.49 , 2.42)	6.84 (1.22 , 22.52)

Days	$\widetilde{\beta}_{t2}$	$\widetilde{\beta}_{t3}$	$\widetilde{\beta}_{t4}$
300	-0.25 (-2.01 , 0.17)	0.22 (-0.55 , 1.50)	-0.31 (-1.51 , 4.50)
500	-0.16 (-1.98 , 0.46)	0.27 (-0.20 , 2.16)	-0.33 (-1.45 , 4.11)
1000	-0.42 (-2.54 , -0.04)	0.29 (-0.26 , 2.00)	0.36 (-0.84 , 6.81)
1500	-0.76 (-4.50 , -0.08)	0.15 (-0.53 , 2.07)	3.11 (-0.12 , 16.09)
2000	-1.51 (-4.89 , -0.35)	1.17 (0.09 , 3.13)	3.69 (-0.51 , 11.73)
2500	-2.60 (-6.92 , -0.52)	1.34 (-0.26 , 3.52)	8.99 (1.04 , 25.26)
3000	-2.87 (-7.19 , -1.03)	1.49 (-0.08 , 3.69)	9.78 (1.98 , 26.05)
3500	-1.14 (-5.36 , -0.18)	0.20 (-1.01 , 2.15)	3.84 (-0.51 , 17.94)

Table 5.18: Estimates of $\theta_t^{(\mathrm{CRP})}$, $\theta_t^{(\mathrm{SAA})}$, $\theta_t^{(\mathrm{IL\text{-}6})}$, $\theta_t^{(\mathrm{tHcy})}$, and $\theta_t(\beta_t)$ at the selected time points

Days	$\widehat{\theta}_t^{(\mathrm{CRP})}$	$\widehat{\theta}_t^{(\mathrm{SAA})}$	$\widehat{\theta}_t^{(\mathrm{IL\text{-}6})}$	$\widehat{\theta}_t^{(\mathrm{tHcy})}$	$\widehat{\theta}_t(\widehat{\beta}_t)$	$\widehat{\theta}_t(\bar{\beta}_t)$
300	0.638	0.629	0.653	0.730	0.777	0.738
500	0.616	0.645	0.653	0.734	0.764	0.732
1000	0.598	0.618	0.666	0.719	0.752	0.745
1500	0.605	0.592	0.681	0.683	0.737	0.732
2000	0.596	0.620	0.668	0.666	0.733	0.725
2500	0.596	0.609	0.663	0.627	0.697	0.684
3000	0.596	0.623	0.678	0.634	0.707	0.695
3500	0.576	0.589	0.634	0.642	0.680	0.675

Days	$\widetilde{\theta}_t^{(\mathrm{CRP})}$	$\widetilde{\theta}_t^{(\mathrm{SAA})}$	$\widetilde{\theta}_t^{(\mathrm{IL\text{-}6})}$	$\widetilde{\theta}_t^{(\mathrm{tHcy})}$	$\widetilde{\theta}_t(\widetilde{\beta}_t)$	$\widetilde{\theta}_t(\bar{\beta}_t)$
300	0.640	0.641	0.663	0.730	0.779	0.739
500	0.623	0.652	0.659	0.733	0.768	0.735
1000	0.606	0.627	0.672	0.724	0.760	0.751
1500	0.611	0.601	0.687	0.690	0.744	0.742
2000	0.595	0.626	0.673	0.673	0.745	0.737
2500	0.598	0.616	0.672	0.633	0.708	0.697
3000	0.599	0.630	0.687	0.636	0.724	0.711
3500	0.523	0.550	0.597	0.635	0.675	0.668

Chapter 6

Discussion

6.1 Concluding Remarks

Based on censored survival data, a unified approach to make inference about the time-dependent AUC and PAUC is developed. We express these accuracy measures as a functional of the joint survival function $S(t, y)$ and simplify the estimation problem. Under different types of censoring mechanisms, two easily computed nonparametric estimators are proposed with rigorously established asymptotic Gaussian processes. Together with the directly estimated variance-covariance functions and the asymptotic i.i.d. representations, statistical inferences about the time-dependent AUC and PAUC are constructed. Provided $S(t, y)$ is estimable, our estimation procedures can also be successfully applied to data with various truncation and/or censoring mechanisms. As mentioned in the paper of Cai et al. (2006), the cases and controls considered for the time-dependent ROC analysis are more suitably defined as $\{T = t\}$ and $\{T > \tau\}$ (or $\{T > t\}$), respectively, in some applications such as the breast cancer study, where τ is a pre-specified time point or the end of study. The proposed methods can be naturally extended to these settings with a slight modification. This extension is reasonable at least in the case with discrete failure time.

Since the performance of a biomarker might be influenced by possible risk factors, an appropriate regression model for the time-dependent AUC is usually used to

delve into their relationship. Another important issue arises when multiple biomarkers are available as we have encountered in the Angiography CAD Study. It is desire to seek combinations of CRP, SAA, IL-6, and tHcy to achieve higher classification abilities at different time points. Our proposed nonparametric estimation method for the time-dependent AUC can be applied to these two topics. For the regression model of the time-dependent AUC, we are motivated by spirit of the estimator $\widetilde{\theta}_t(\widetilde{q}_{1t})$ to develop estimating equations for the parameters of interest. Under the validity of a time-varying coefficient EGLM with unspecified link function $G_t(\cdot)$, the nonparametric estimation procedures for the optimal composite biomarkers are obtained via maximizing the modified time-dependent AUC quantities. The predictor $\widehat{\beta}_t^T Y$ is found to be appropriate in predicting the vital statuses of subjects in the numerical studies, while the predictor $\widetilde{\beta}_t^T Y$ is shown to be suitable even if the censoring time is highly correlated with the baseline markers. We also provide estimation methods for the corresponding accuracy measures $ROC_t(\beta_t)$ and $\theta_t(\beta_t)$ of $\beta_t^T Y$. As we can see the time-varying coefficient EGLM is very flexible and contains many practical used models as special cases such as the Cox's proportional hazards model, the GAFT model, and the time-varying coefficient logistic regression model. Our methods can also be used to diagnose the appropriateness of model specification. The rationale behind this fact is that the estimated time-dependent AUC based on a specific model should be close to that based on our nonparametric estimators if the model is correctly specified.

When the assumption of marker-dependent censorship is valid, it is usually unavoidable to use smoothing techniques which will complicate the development of the corresponding inference. Obviously, this can be seen in the issue of combining biomarkers where a higher order kernel function is required to ensure $\sqrt{n}$-consistency. When the number of biomarkers is large, the curse of dimensionality will be encountered and the estimation will become more unstable. It is evidenced by our limited simulation studies that the estimators proposed under totally independent censorship are robust toward the violation of marker-dependent censorship. In applications, they

have the advantage that no smoothing technique is involved. Thus, we suggest to use and compare both methods even the behavior of censoring is believed to be highly correlated with biomarkers.

6.2 Future Works

6.2.1 Multiple Biomarker Comparison

The time-dependent ROC curve as well as its summary measures are usually used to evaluate the classification abilities of biomarkers. As shown in the analysis of ACTG 175 study in Chapter 5, our proposed methodologies can be extended to compare the corresponding time-dependent AUCs and PAUCs of several biomarkers. On the other hand, different accuracy measures might be more preferable in the assessment. Provided a subject being diagnose to be disease at the current stage, a scientific question might be raised concerning the probability that the subject is really diseased in the future. Thus, the prediction ability is more relevant to the end user or subject being diagnosed. Two measure indices are widely applied in the assessment: the time-dependent positive predictive value $PPV_t(q) = P(T \leq t | F_Y(Y) > q)$ and the time-dependent negative predictive value $NPV_t(q) = P(T > t | F_Y(Y) \leq q)$ with $F_Y(y) = P(Y \leq y)$ (Zheng et al. (2008)), display the prediction accuracy at various quantile value q. One can observe that the higher the curves, the better the biomarker in prediction. The given condition $\{F_Y(Y) > q\}$ and $\{F_Y(Y) \leq q\}$ represent "predict $100(1-q)\%$ of subjects to be diseased" and "predict $100q\%$ of subjects to be non-diseased", respectively. The purpose of using $F_Y(y)$ in the definitions is mainly to facilitate biomarkers with different measurement scales to have the same base of comparison, i.e., the same q. Interestingly, we can derive that

$$PPV_t(q) = 1 - \frac{S_T(t) - (1-q) \cdot NPV_t(q)}{q}, \tag{6.1}$$

which is a monotone function of $NPV_t(q)$ for any fixed q. It's further implied from this fact that we only need to consider $PPV_t(q)$ in the comparison of prediction

abilities. Same with the role of the AUC in the time-dependent ROC curve analysis, the quantity $\int_0^1 PPV_t(q)dq$ can be used as a summary measure of prediction ability.

Let the biomarker $Y^{(k)}$ have the corresponding $TPR_t^{(k)}(y)$, $FPR_t^{(k)}(y)$, and $PPV_t^{(k)}(q)$, $k = 1, 2$. Recall that we adopt that measure $\gamma_\alpha(t)$ in 5.2.1 to compare the classification abilities. As to the case of assessing prediction abilities, the quantity

$$\mu(t) = \int_0^1 \{PPV_t^{(1)}(q) - PPV_t^{(2)}(q)\}dq \tag{6.2}$$

provides useful information. The positive (negative) value of $\gamma_\alpha(t)$ or $\mu(t)$ then indicate the superiority (inferiority) of $Y^{(1)}$ is classification or prediction. Just like we have discussed the insensitivity of the AUC in Chapter 2, two total different biomarkers may result in $\mu(t) = 0$. The more sensitive summary measures of the difference between $Y^{(1)}$ and $Y^{(2)}$ are the area between the curves (ABC), which are defined as

$$ABC_{ROC}(t) = \int_{-\infty}^{\infty} |TPR_t^{(1)}(FPR_t^{(1)-1}(u)) - TPR_t^{(2)}(FPR_t^{(2)-1}(u))|du \tag{6.3}$$

and

$$ABC_{PPV}(t) = \int_0^1 |PPV_t^{(1)}(q) - PPV_t^{(2)}(q)|dq. \tag{6.4}$$

The magnitude of difference between $Y^{(1)}$ and $Y^{(2)}$ in classification (prediction) abilities is then reflected by the distance of $ABC_{ROC}(t)$ ($ABC_{PPV}(t)$) and zero.

In some empirical examples such as the paired-design experiment in clinical trial where each subject undergoes both diagnostic tests, data of the form $\{X_i, \delta_i, Y_i^{(1)}, Y_i^{(2)}\}_{i=1}^n$ are often collected. The scientific interests usually focus on the comparison of the performances between $Y^{(1)}$ and $Y^{(2)}$. We can re-expressed the aforementioned summary measures as functionals of $S^{(1)}(t, y) = P(T > t, Y^{(1)} > y)$ and $S^{(2)}(t, y) = P(T > t, Y^{(2)} > y)$. Under the totally independent censoring (C and $(T, Y^{(1)}, Y^{(2)})$ are independent), the proposed estimation criterion can be naturally applied to make statistical inferences about these quantities directly because $S^{(1)}(t, y)$ and $S^{(2)}(t, y)$ can be marginally estimated by $\{X_i, \delta_i, Y_i^{(1)}\}_{i=1}^n$ and $\{X_i, \delta_i, Y_i^{(2)}\}_{i=1}^n$. As an illustrative example in the CAD study, the censoring mechanism might relate to the baseline

biomarkers. The assumptions of marker-dependent censorship made separately on $(T, C, Y^{(1)})$ and $(T, C, Y^{(2)})$ are clearly inappropriate. Under a more flexible marker-dependent censoring assumption (T and C are independent conditioning on $Y^{(1)}$ and $Y^{(2)}$), the estimators used for $S^{(1)}(t, y)$ and $S^{(2)}(t, y)$ in this article need to be further modified. To circumvent this difficulty, Akritas (1994) suggested using

$$\widehat{S}(t, y_1, y_2) = \frac{1}{n} \sum_{i=1}^{n} \widehat{S}(t|Y_i^{(1)}, Y_i^{(2)}) I(Y_i^{(1)} > y_1, Y_i^{(2)} > y_2), \tag{6.5}$$

to estimate $S(t, y_1, y_2) = P(T > t, Y^{(1)} > y_1, Y^{(2)} > y_2)$, where $\widehat{S}(t|y_1, y_2)$ is a smoothing estimator of $P(T > t|Y^{(1)} = y_1, Y^{(2)} = y_2)$ (cf. Beran (1981)). Thus, $S^{(1)}(t, y)$ and $S^{(2)}(t, y)$ can be separately estimated by $\widehat{S}(t, y, -\infty)$ and $\widehat{S}(t, -\infty, y)$. It is worthwhile to investigate the related comparison procedures in the future study.

6.2.2 Optimality in Classification and Prediction

It is found in our recent research that the problem of prediction power can be equivalently transformed into that of classification ability. More precisely speaking, we obtain that a biomarker with the highest time-dependent ROC curve will possess the largest $PPV_t(q)$ and, hence, $NPV_t(q)$ for any q, and vice versa. One interesting topic arises in whether we can find another estimation procedure for the parameter of the optimal composite biomarker. Instead of maximizing the time-dependent AUC, we seek the optimal markers that maximizes the area under the curve of $PPV_t(q)$. Although it can be concluded that these criteria are the same from the above arguments, the obtained estimators of optimal composite biomarkers might be different. The gain and loss of these estimators will be promised to be deeply studied in future.

As we have discussed in Chapter 2, the AUC is an overall evaluation and some useful information may not be captured. To significantly improve the classification accuracy, an alternative strategy is to find an optimal combination in the sense that the time-dependent true positive rate is maximized at the same value of false positive rate. The optimization can be also achieved via maximizing the time-dependent PAUC of composite biomarker. According to the research purposes, our proposed

estimators for the time-dependent PAUC, $TPR_t(c, \beta)$, and $FPR_t(c, \beta)$ can be naturally generalized to this issue and should be investigated.

6.2.3 Partly Conditional Time-Dependent AUC

Throughout this thesis, we only focus on the classification ability of a baseline biomarker for time-dependent vital status. When the considered biomarker has an underlying continuous-time stochastic process $\{Y(s) : s \in [0, \tau]\}$, the succeeding measured biomarker might be more predictive than the baseline one. Thus, we can consider a more flexible partly conditional partly conditional AUC to characterize the updated information in classifying the binary vital status, which is formulated as

$$\theta_{st} = P(Y_i(s) > Y_j(s) | s < T_i \leq t, T_j > t) \text{ for } i \neq j, 0 < s < t.$$

In some applications such as the ACTG 175 study, a test result might be intermittently collected at multiple follow-up times. Based on the censored survival data $\{X_i, \delta_i, Y_i(s_{i1}), \cdots, Y_i(s_{in_i})\}_{i=1}^n$ with s_{ij}'s being longitudinal measurement times, the corresponding statistical inference procedure for θ_{st} remains to be established in future studies. Again, when the time-dependent or time-independent covariates $Z(s)$ are considered as possible risk factors on the performance of $Y(s)$, the covariate-specific time-dependent AUC

$$\theta_{st}(Z_i, Z_j) = P(Y_i(s) > Y_j(s) | s < T_i \leq t, T_j > t, Z_i(s), Z_j(s)) \text{ for } 0 < s < t,$$

can be appropriately modeled by the the time-varying coefficient regression model

$$\theta_{st}(Z_i, Z_j) = h(\gamma_{st} Z_{ij}(s)),$$

where h is a response function and $Z_{ij}(s)$ is a function of $Z_i(s)$ and $Z_j(s)$. Similar to the discussion in Chapter 3, the estimating equations for γ_{st} can be constructed from the estimation method for θ_{st}.

Bibliography

[1] Akritas, M. G. (1994). Nearest neighbor estimation of a bivariate distribution under random censoring. *Annals of Statistics*, 22, 1299-1327.

[2] Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical report, Univ. California, Berkeley.

[3] Burke, M. D. (1988). Estimation of a bivariate distribution function under random censorship. *Biometrika*, 75, 379-382.

[4] Cai, T. and Cheng, S. (2008). Robust combination of multiple diagnostic tests for classifying censored event times. *Biostatistics*, 9, 216-233.

[5] Cai, T. and Dodd, L. E. (2008). Regression analysis for the partial area under the ROC curve. *Statistica Sinica*, 18, 817-836.

[6] Cai T., Pepe M. S., Zheng Y, Lumley, T., and Henny, N. S. (2006). The sensitivity and specificity of markers for event times. *Biostatistics*, 7, 182-197.

[7] Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika*, 68, 417-423.

[8] Chambless, L. E. and Diao, G. (2006). Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Statistics in Medicine*, 25, 3474-3486.

[9] Chiang, C. T. and Huang, S. Y. (2009). Estimation for the optimal combination of markers without modeling the censoring distribution. *Biometrics*, 65, 152-158.

[10] Chiang, C. T., Wang, S. H., and Hung, H. (2008). Random weighting and Edgeworth expansion for the nonparametric time-dependent AUC estimator. *Statistica Sinica*. To appear.

[11] Dabrowska, D. M. (1987). Uniform consistency of nearest neighbor and kernel conditional Kaplan-Meier estimates. Technical Report No. 86, Univ. California, Berkeley.

[12] Dabrowska, D. M. (1989). Uniform Consistency of the Kernel Conditional Kaplan-Meier Estimate. *Annals of Statistics*, 17, 1157-1167.

[13] Daouia, A., Florens, J. P., and Simar, L. (2008). Functional convergence of quantile-type frontiers with application to parametric approximations. *Journal of Statistical Planning and Inference*, 138, 708-725.

[14] Dodd, L. E. and Pepe, M. S. (2003a). Semiparametric regression for the area under the receiver operating characteristic curve. *Journal of the American Statistical Association*, 98, 409-417.

[15] Dodd, L. E. and Pepe, M. S. (2003b). Partial AUC estimation and regression. *Biometrics*, 59, 614-623.

[16] Du, Y. and Akritas, M. G. (2002). Uniform strong representation of the conditional Kaplan-Meier process. *Mathematics Methods of Statistics*, 11, 152-182.

[17] Dwyer, A. J. (1996). In pursuit of a piece of the ROC. *Radiology*, 201, 621-625.

[18] Emir, B., Wieand, S., Jung, S. H., and Ying, Z. (2000). Comparison of diagnostic markers with repeated measurements: a non-parametric ROC curve approach. *Statistics in Medicine*, 19, 511-523.

[19] Fleming, T. R. and Harrington, D. P. (1991). Counting process and survival analysis. New York: Wiley.

[20] Hammer, S. M., Katzenstein, D. A., Hughes, M. D., Gundacker, H., Schooley, R. T., Haubrich, R. H., Henry, W. K., Lederman, M. M., Phair, J. P., Niu, M., Hirsch, M. S., and Merigan, T. C. (1996). A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *The New England Journal of Medicine*, 335, 1081-1090.

[21] Heagerty, P. J., Lumley, T., and Pepe, M. (2000). Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*, 56, 337-344.

[22] Horowitz, J. L. (1992). A smoothed maximum score estimator for the binary response model. *Econometrica*, 60, 505-531.

[23] Horvitz, D. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663-685.

[24] Jiang, Y., Metz, C. E., and Nishikawa, R. M. (1996). A receiver operating characteristic partial area index for highly sensitive diagnostic tests. *Radiology*, 201, 745-750.

[25] Khan, S. and Tamer, E. (2007). Partial rank estimation of duration models with general forms of censoring. *Journal of Econometrics*, 136, 251-280.

[26] Koul, H., Susarla, V., and Van Ryzin, J. (1981). Regression analysis with randomly right censored data. *Annals of Statistics*, 9, 1276-1288.

[27] Lee, K. W. J., Hill, J. S., Walley, K. R., and Frohlich, J. J. (2006). Relative value of multiple plasma biomarkers as risk factors for coronary artery disease and death in an angiography cohort. *Canadian Medical Association Journal*, 174, 461-466.

[28] Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. L. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society Series B*, 62, 711-730.

[29] McClish, D. K. (1989). Analyzing a portion of the ROC curve. *Medical Decision Making*, 9, 190-195.

[30] Metz, C. E., Herman, B.A., and Shen, J. (1998). Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuous distributed data. *Medical in Medicine*, 17, 1033-1053.

[31] Newey, W. K. and McFadden, D. (1994). Large sample estimation and hypothesis testing. In: McFadden, D. and Engler, R. (editors), *Handbook of Econometrics*, Volume IV. North Holland: Amsterdam, 2113-2245.

[32] Nolan, D. and Pollard, D. (1987). U-processes: Rates of convergence. *Annals of Statistics*, 15, 780-799.

[33] Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 57, 1027-1057.

[34] Pepe, M. S. (1997). A regression modelling framework for receiver operating characteristic curves in medical diagnostic testing. *Biometrika*, 84, 595-608.

[35] Pepe, M. S. (1998) Three approaches to regression analysis of receiver operating characteristic curves for continuous test results. *Biometrics*, 54, 124-135.

[36] Pepe, M. S. (2000) Receiver operating characteristic methodology. *Journal of the American Statistical Association*, 95, 307-311.

[37] Pepe, M. S. (2003) The statistical evaluation of medical tests for classification and prediction. Oxford, New York.

[38] Pollard, D. (1990). Empirical processes: Theory and applications. Hayward: Institute of Mathematical Statistics.

[39] Robins, J. M. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell N, Dietz K, Farewell V (ed) AIDS Epidemiology-Methodological Issues, Birkhäuser, Boston, 297-331

[40] Serfling, R. J. (1980). Approximation theorems of mathematical statistics. New York: Wiley.

[41] Sherman, R. P. (1993). The limiting distribution of the maximum rank correlation estimator. *Econometrica*, 61, 123-137.

[42] Sherman, R. P. (1994a). Maximal inequalities for degenerate U-processes with applications to optimization estimators. *Annals of Statistics*, 22, 439-459.

[43] Sherman, R. P. (1994b). U-process in the analysis of a generalized semiparametric regression estimator. *Econometric Theory*, 10, 372-395.

[44] Song, X., Ma, S., Huang, J., and Zhou, X. H. (2007). A semiparametric approach for the nonparametric transformation survival model with multiple covariates. *Biostatistics*, 8, 197-211.

[45] van der Vaart, A. W. (2000). Asymptotic Statistics. Cambridge: Cambridge University Press.

[46] Zhang, D. D., Zhou, X. H., Freeman, Daniel H. J., and Freeman, J. L. (2002). A non-parametric method for the comparison of partial areas under ROC curves and its application to large health care data set. *Statistics in Medicine*, 21, 701-715.

[47] Zheng, Y., Cai, T., and Feng, Z. (2006). Application of the time-dependent ROC curves for prognostic accuracy with multiple biomarkers. *Biometrics*, 62, 279-287.

[48] Zheng, Y., Cai, T., Pepe, M. S., and Levy W. C. (2008). Time-dependent predictive value of prognostic biomarkers with failure time outcome. *Journal of the American Statistical Association*, 103, 362-368.

[49] Zhou, X. H., McClish, D. K., and Obuchowski, N. A. (2002). Statistical methods in diagnostic medicine. New York: Wiley.

[50] Zou, K. H., Hall, W. J., and Shapiro, D. E. (1997). Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests. *Statistics in Medicine*, 16, 2143-2156.