4-21-2008

Direct Adjustment Method on Aalen's Additive Hazards Model for Competing Risks Data

Haci Mustafa Akcin

DIRECT ADJUSTMENT METHOD ON AALEN'S ADDITIVE HAZARDS MODEL FOR COMPETING RISKS DATA

by

HACI MUSTAFA AKCIN

Under the Direction of Dr. Xu Zhang

ABSTRACT

Aalen's additive hazards model has gained increasing attention in recently years because it model all covariate effects as time-varying. In this thesis, our goal is to explore the application of Aalen's model in assessing treatment effect at a given time point with varying covariate effects. First, based on Aalen's model, we utilize the direct adjustment method to obtain the adjusted survival of a treatment and comparing two direct adjusted survivals, with univariate survival data. Second, we focus on application of Aalen's model in the setting of competing risks data, to assess treatment effect on a particular type of failure. The direct adjusted cumulative incidence curve is introduced. We further construct the confidence interval of the difference between two direct adjusted cumulative incidences, to compare two treatments on one risk.

INDEX WORDS: Direct adjustment method, Aalen's additive hazards model, Competing risks, Survival analysis, Cumulative incidence function, Survival function

DIRECT ADJUSTMENT METHOD ON AALEN'S ADDITIVE HAZARDS MODEL FOR COMPETING RISKS DATA

by

HACI MUSTAFA AKCIN

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

Master of Science

in the College of Arts and Sciences

Georgia State University

2008

DIRECT ADJUSTMENT METHOD ON AALEN'S ADDITIVE HAZARDS MODEL FOR COMPETING RISKS DATA

by

HACI MUSTAFA AKCIN

Committee Chair: Dr.Xu Zhang

Committee: Dr.Yu-Sheng Hsu

Dr.Yichuan Zhao

Dr.Jiawei Liu

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2008

DEDICATION

I would like to dedicate this thesis to all of the great minds of Statistics, who worked restlessly to find best solutions to make science and human life to improve. I admire you.

ACKNOWLEDGEMENTS

I would like to acknowledge the help and support of those who helped and supported me. First of all, I am very thankful to my advisor Dr.Zhang who had been a great advisor, a patient and a dedicated professor. This thesis cannot be finished without her supervision. I have taken all of her classes and they have been a great learning experience. I would also like to acknowledge the other members of my committee, Dr. Yu-Sheng Hsu, Dr.Yichuan Zhao, and Dr. Jiawei Liu for being very kind to be part of my committee and taking the time to read this thesis and provide very useful comments. I'd also like to thank all other professors in the Mathematics & Statistics Department at Georgia State University where I felt very comfortable and welcomed by every one of them when I needed any help or assistance. I thank you all for everything. Finally, I want to acknowledge the love of my life, my best friend and my soul-mate, my wife Simona-Daniela Dima for her support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BCS: Breast Conserving Surgery

k-score: Karnofsky score

CHAPTER I. INTRODUCTION

The survival curves of different treatments are often presented in medical articles to visualize the efficacy of treatments. When the failure time data is obtained from a clinical trial, one can directly present the crude survival curve of each treatment group. When observational data is utilized, one needs to analyze data with appropriate regression model and compute the adjusted survival curve of a treatment based on the regression model.

There are two methods to compute the adjusted survival curve of a treatment based on a regression model. The first method is known as the mean covariate method. For this method, one computes the survival probabilities for an individual with the average covariate values, given a treatment. Let $\bar{\boldsymbol{Z}}$ be the vector of average covariates. Let A denote treatment assignment, where $A = 1, \cdots, K$. $S(t; A = k, \bar{\boldsymbol{Z}})$ should be computed and viewed as the survival probability for treatment k. This method has the drawback that the covariate value for the average patient may be meaningless for categorical variables. For example, if the covariate gender is coded as 0 for male and 1 for female, it would be difficult to interpret the survival for an average patient the gender covariate value 0.7.

The second method is the direct adjustment method, which computes the average survival curve over the entire sample, given a treatment (Makuch, 1982; Chang, Gelman and Pagano, 1982; Gail and Byar, 1986). This method is also known as the group prognosis method. For a sample of size n, let $\boldsymbol{Z}_i$ be the vector of characteristics for the ith individual. The direct adjusted survival curve is given by

$$S_k(t) = \frac{1}{n}\sum_{i=1}^{n} S(t;\ A = k, \mathbf{Z}_i)$$

Lee et al. (1992) and Ghali et al. (1986) provided the computer programs in SAS, STATA, and S-plus for obtaining the direct adjusted survival curves based on a Cox model (Cox,1972). The asymptotic results of the direct adjusted survival curves based on a Cox model and a Weibull regression model were studied by Gail and Byar (1986). Zhang et al. (2007) gave the asymptotic results of such curves based on a stratified Cox model and provided a SAS macro to derive the curves and their confidence limits.

Cox model is the most commonly used regression model to assess covariate effect with failure time data. It is known that Cox model is a multiplicative hazards model, and the effect of covariate is multiplicative on the baseline level. Another principle framework to analyze failure time data is additive models, in which the hazard of a risk factor is often an additive increment on the baseline hazard. A nonparametric additive hazards model was proposed by Aalen (1989, 1993). Aalen gave the least-square estimation of the regression coefficients. Huffer and McKeague (1991) proposed a weighted least-square estimation approach, which specifically addresses the problem of the unequal variances. Lin and Ying (1994) proposed an additive hazards model that assumes constant covariate effect. An additive hazards model that includes both time-varying effect and constant effect was studied by McKeague and Sasieni (1994).

Aalen's model has gained important attention in recent years because the effects of all covariates are modeled as time-varying (Martinussen and Scheike, 2006). Such a model does not require the constant proportional or additive hazards assumption between levels of a covariate, which is crucial for other regression models like Cox model or Lin and Ying's model. In this thesis, we wish to explore the application of Aalen's model in assessing treatment effect with

time-varying covariate effect. The aforementioned direct adjustment method shall be adopted to give adjusted treatment effect. Our first goal is to estimate direct adjusted survival of a treatment, and compare the direct adjusted survivals of two treatments at a given time. This has been implemented in a SAS macro and illustrated in stem cell transplant example.

The second goal of this thesis is to illustrate the application of Aalen's model with competing risks data. It has been pointed out that additive hazard model is consistent with the nature of competing risks data (Klein, 2005). It often appears in clinical articles that, with competing risks data, all cause-specific hazards and the all-cause hazard are specified as Cox models. Klein explicitly criticized this approach. Consider a study with two competing risks. Let $\lambda_1(t;\mathbf{z})$ and $\lambda_2(t;\mathbf{z})$ be cause-specific hazards. Let $\lambda_\bullet(t;\mathbf{z})$ be the hazard for the all-cause failure time. By definition, $\lambda_\bullet(t;\mathbf{z}) = \lambda_{01}(t;\ \mathbf{z}) + \lambda_{02}(t;\ \mathbf{z})$. If one wishes to model all the above hazard rate functions by Cox models, that is,

$$\lambda_1(t;\mathbf{z}) = \lambda_{01}(t)\exp(P_1^T\mathbf{z}),$$
$$\lambda_2(t;\mathbf{z}) = \lambda_{02}(t)\exp(P_2^T\mathbf{z}),$$
$$\lambda_\bullet(t;\mathbf{z}) = \lambda_{0\bullet}(t)\exp(P_\bullet^T\mathbf{z}),$$

then one would expect the following relationship,

$$\lambda_{0\bullet}(t)\exp(P_\bullet^T\mathbf{z}) = \lambda_{01}(t)\exp(P_1^T\ \mathbf{z}) + \lambda_{02}(t)\exp(P_2^T\ \mathbf{z}).$$

Klein explained that the above equation holds only in some rare situations and the problem of internal inconsistency exists. Klein suggested additive hazard model for competing risks problems and specifically analyzed a real example with Aalen's model as well as Lin and Ying's model. In Klein's paper, only the individual covariate adjusted cause-specific hazard and cumulative incidence function were presented.

In this thesis we apply the direct adjustment method to compute the cause-specific hazard and cumulative incidence function, on the basis of Aalen's additive model. We give the confidence limits for the direct adjusted cause-specific hazard and cumulative incidence function for one type of failure. Treatment comparison regarding a given cause is implemented by constructing the confidence intervals of the difference between two direct adjusted cumulative incidence functions. We apply the methods to data of breast cancer patients from SEER registry. In this example, three competing risks exist: breast cancer, vascular disease, and other causes. The primary goal is to compare two treatments: breast conserving surgery and mastectomy. We wish to assess the effectiveness of treatments with respect to each type of failure and the all-cause failure.

The thesis is organized in the following way. Chapter 2 consists of three sections. In Section 2.1, we briefly review Aalen's model. In Section 2.2, we give the asymptotic results for the direct adjusted survival curve based on Aalen's model, and then describe the macro that reports the direct adjusted survival curves. In Section 2.3 we utilize the Aalen's model to analyze competing risks data; we illustrate how to compare two treatments for a particular type of failure at a given time. The results from a simulation study are reported in Chapter 3. We analyze two real examples in Chapter 4 to illustrate the methods described in this thesis. The concluding remarks are given in Chapter 5.

CHAPTER II. METHODS

2.1 The Additive Hazards Model

Let *T* and *C* denote the failure time and the censoring time. For a sample of size *n*, the observed data can be summarized as $\{\widetilde{T}_i, \Delta_i, \boldsymbol{Z}_i\}$ for $i = 1, \cdots, n$, where $\widetilde{T}_i = \min(T_i, C_i)$ and $\Delta_i = I(T_i \le C_i)$, $\boldsymbol{Z}_i$ is a *p*-dimensional vector of covariates with $Z_i = (Z_{i1}(t), \cdots, Z_{ip}(t))$. Note that we have set the covariate vector as a row vector. Let $N_i(t)$ be the counting process for the *i*th subject of the sample, where $N_i(t) = I(\widetilde{T}_i \le t, \Delta_i = 1)$. Aalen's model specify the intensity process of $N_i(t)$ as

$$\lambda_i(t) = Y_i(t)\big(\beta_0(t) + \beta_1(t)Z_{i1}(t) + \cdots + \beta_p(t)Z_{ip}(t)\big), \qquad 2.1$$

where $Y_i(t) = I(\widetilde{T}_i \ge t)$, is the at risk indicator, and $\boldsymbol{\beta}(t) = (\beta_0(t), \beta_1(t), \cdots, \beta_p(t))^T$ is a vector of regression coefficients. The integrated regression coefficient is given by

$$\boldsymbol{B}(t) = \int_0^t \boldsymbol{\beta}(u)du\,. \qquad 2.2$$

Aalen (1989) considered least-square estimation of $\boldsymbol{B}(t)$. Let $\boldsymbol{N}(t) = (N_i(t); i = 1, \cdots, n)$ and let $\boldsymbol{Y}(t)$ be a $n \times (p+1)$ matrix with the *i*th row given by

$$Y_i(t)(1, Z_{i1}(t), \cdots, Z_{ip}(t))\,. \qquad 2.3$$

The Nelson-Aalen estimator of $\boldsymbol{B}(t)$ is given by

$$\hat{\boldsymbol{B}}(t) = \int_0^t J(u)\boldsymbol{Y}^-(u)\,\mathrm{d}\,\boldsymbol{N}(u) \qquad 2.4$$

where

$$\boldsymbol{Y}^-(t) = (\boldsymbol{Y}(t)^T\boldsymbol{Y}(t))^{-1}\boldsymbol{Y}(t)^T\,. \qquad 2.5$$

and $J(u) = I(\text{rank of } Y(u) = p+1)$.

Define

$$\boldsymbol{B}^*(t) = \int_0^t J(u)\,\boldsymbol{\beta}(u)du\,. \qquad 2.6$$

An estimator of the mean squared error function,

$$\boldsymbol{\Sigma}(t) = \boldsymbol{E}\,(\,\hat{\boldsymbol{B}} - \boldsymbol{B}^*)(\hat{\boldsymbol{B}} - \boldsymbol{B}^*)^T(t)\,, \qquad 2.7$$

is given by

$$\hat{\boldsymbol{\Sigma}}(t) = \int_0^t J(u)\boldsymbol{Y}^{-}(u)\ \mathrm{diag}\ \mathrm{d}\,\boldsymbol{N}(u)(\boldsymbol{Y}^{-}(u))^T \qquad 2.8$$

It is often desirable to predict the survival probability for an individual with given covariates. Given $\boldsymbol{z}$, the cumulative hazard at t, $A(t;\boldsymbol{z})$, can be estimated by

$$\hat{A}(t;\mathrm{z}) = \int_0^t \boldsymbol{z}\ \ \mathrm{d}\,\hat{\boldsymbol{B}}(u)\,. \qquad 2.9$$

An estimator of survival probability is

$$\hat{S}(t;\boldsymbol{z}) = \exp(-\hat{A}(t;\boldsymbol{z}))\,. \qquad 2.10$$

For fixed $\boldsymbol{z}$, variance of $\hat{A}(t;\boldsymbol{z})$ is given by $\boldsymbol{z}\boldsymbol{\Sigma}(t)\boldsymbol{z}^T$, which can be estimated by $\boldsymbol{z}\hat{\boldsymbol{\Sigma}}(t)\boldsymbol{z}^T$

2.2 The Direct Adjusted Survival Curves and the Macro

Suppose that a study involves m treatments. We wish to estimate the overall survival for the ith treatment for $i = 1, \ldots, m$. We can split the covariate vector $\boldsymbol{Z}$ in two components, $\widetilde{\boldsymbol{X}}$ and $\widetilde{\boldsymbol{Z}}$, where $\widetilde{\boldsymbol{X}}$ is the coding for treatment assignment and $\widetilde{\boldsymbol{Z}}$ specifies patient's characteristics. Specifically, $\boldsymbol{Z} = (\,\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{Z}}\,)$. Please note that $\widetilde{\boldsymbol{X}}$ is fixed and $\widetilde{\boldsymbol{Z}}$ can be time-dependent. We further let X be the treatment index, corresponding to a specific coding of $\widetilde{\boldsymbol{X}}$. For example, consider a study involving four treatments, we will let $X \in \{1, 2, 3, 4\}$ and treat the effect of treatment 1 as

baseline; $X = 1$ would be equivalent to $\widetilde{\boldsymbol{X}} = (0, 0, 0)$, $X = 2$ would be equivalent to $\widetilde{\boldsymbol{X}} = (1, 0, 0)$, and etc. We define

$$\boldsymbol{z}^i = (1, X = i, \widetilde{\boldsymbol{Z}} = \boldsymbol{z}), \qquad i = 1, \ldots, m, \tag{2.12}$$

with the understanding that the value of X should be replaced by the corresponding value of $\widetilde{\boldsymbol{X}}$. We further add a subscript l to $\boldsymbol{z}^i$ to let $\widetilde{\boldsymbol{Z}}$ be the characteristics of the lth patient. Let $S_i(t)$ denote the direct adjusted survival of the ith treatment for $i = 1, \ldots, m$, that is, $S_i(t) = n^{-1}\sum_l S(t; \boldsymbol{z}_l^i)$. $S_i(t)$ can be estimated by

$$\hat{S}_i(t) = \frac{1}{n}\sum_{l=1}^{n} \exp(-\hat{A}(t; \boldsymbol{z}_l^i)) \tag{2.13}$$

Applying the delta method, we derive the variance of $S_i(t)$, which can be estimated by

$$\hat{\text{var}}[\hat{S}_i(t)] = \left(n^{-1}\sum_{l=1}^{n} \hat{S}(t; \boldsymbol{z}_l^i)\boldsymbol{z}_l^i\right)\hat{\boldsymbol{\Sigma}}(t)\left(n^{-1}\sum_{l=1}^{n} \hat{S}(t; \boldsymbol{z}_l^i)\boldsymbol{z}_l^i\right)^T \tag{2.14}$$

It is important to study the difference between two direct adjusted survivals,

$$\pi_{ij}(t) = S_i(t) - S_j(t), \quad i, j \in \{1, \cdots, K\}, i \neq j. \tag{2.15}$$

It can be estimated by $\hat{\pi}_{ij}(t) = \hat{S}_i(t) - \hat{S}_j(t)$. The variance of $\hat{\pi}_{ij}(t)$ can be estimated by

$$\begin{aligned}\hat{\text{var}}[\hat{\pi}_{ij}(t)] = &\left(n^{-1}\sum_{l=1}^{n} \hat{S}(t; \boldsymbol{z}_l^i)\boldsymbol{z}_l^i - n^{-1}\sum_{l=1}^{n} \hat{S}(t; \boldsymbol{z}_l^j)\boldsymbol{z}_l^j\right) \\ &\hat{\boldsymbol{\Sigma}}(t)\left(n^{-1}\sum_{l=1}^{n} \hat{S}(t; \boldsymbol{z}_l^i)\boldsymbol{z}_l^i - n^{-1}\sum_{l=1}^{n} \hat{S}(t; \boldsymbol{z}_l^j)\boldsymbol{z}_l^j\right)^T\end{aligned} \tag{2.16}$$

We have implemented a SAS macro to compute the above estimators. Before we run the macro we should prepare a SAS data set to contain the following variables: a variable of the

failure time, a censoring indicator (which should be coded as 1 for an event occurs and 0 for censoring), a variable of treatment which should be coded as 1,…*K*, and the variables for patients' characteristics. This SAS macro can be used for both continuous and categorical variables. An *m*-level categorical prognostic risk factor should be coded as (*m-1*) binary variables. For example, suppose that race is included in the regression and it contains four levels of White, Black, Hispanic and Other. We may treat White as the reference level and create three binary variables of Black (1 for Black, 0 otherwise), Hispanic (1 for Hispanic, 0 otherwise), and Other (1 for Other, 0 otherwise), for the race factor.

Suppose that our macro is saved as a SAS file with the file name "**estimate.sas**"**.** One can save a copy of the file in the current working directory, and then use the following SAS statement to load the macro into the current program.

% include 'estimate.sas' ;

The macro will be invoked by running the following statement;

% estimate (indata, time, event, group, option, covlist, outdata)

where;

indata the input SAS data set name

time the variable of failure time

event the variable of event indicator

group the variable of treatment indicator

option the variable of covariate indicator (1 with covariates, 0 without)

covlist a list of all covariates

outdata the SAS output data set name

The results of the macro will be saved in a SAS output data set named "**outdata**" and printed in the output window. The output data set includes the failure time variable, the estimated direct adjusted survivals and their estimated standard errors, as well as the estimated standard errors of the differences between two direct adjusted survivals. $\hat{S}_i(t)$'s and their estimated standard errors are given by "surv1, …, surv*K*, se1, …, se*K*". The estimated standard errors $\hat{\sigma}_{i,j}(t)$ for $1 \le i < j \le K$ are given by "se12, …, se$(K-1)K$".

2.3 The Direct Adjusted Cause-Specific Hazards and Cumulative Incidence Functions

In medical researches, the problems of competing risks often arise. For example, breast cancer patients usually have a relatively long-term post-surgery survival. Such patients may finally die from other causes instead of cancer. The survival of bone marrow transplant patients is another example of competing risks. Some patients die shortly after transplant due to graft-versus-host disease, and the other competing risk is leukemia relapse. For the problems of competing risks, the occurrence of one type of failure precludes other types of failure.

It is known that, in the setting of competing risks, the cumulative incidence function is an important quantity to evaluate the cumulative failure probability due to a particular cause, with existence of other types of failure. In some applications, the cause-specific hazard has been utilized as well as criteria to compare effect of different groups. However, its drawback is obvious: the cause-specific hazard is meaningful in the hypothetical setting that other competing risks are removed. Therefore, in this thesis, we emphasize on the cumulative incidence function. For comparison, the result on cause-specific hazard is provided as well.

With covariates, the cumulative incidence function is defined as a function of all cause-specific hazards, $F_k(t;\boldsymbol{z}) = \int_0^t S(t-;\boldsymbol{z})\partial A_k(t;\boldsymbol{z})$, where $S(t;\boldsymbol{z}) = \exp(-\sum_k A_k(t;\boldsymbol{z}))$ and $A_k(t;\boldsymbol{z})$ is the cumulative hazard due to the *k*th cause for given covariate values **z**. Certain regression models needs to be specified for cause-specific hazards. It has been pointed out that the framework of additive hazards models is a solution to modeling the cause-specific hazards with the competing risks data (Klein, 2005). When effects of covariates are time-varying, Aalen's model is the appropriate model. Let ***z*** be the row vector including 1, for intercept, and the covariate values. One needs to specify,

$$\lambda_k(t;\boldsymbol{z}) = \boldsymbol{z}\,\boldsymbol{\beta}_k(t) \quad , \qquad \forall k.$$

The all-cause hazard should be the sum of all cause-specific hazards. Let $\boldsymbol{\beta}_\bullet(t) = \sum_k \boldsymbol{\beta}_k(t)$. Then, Aalen's model can properly model the all-cause hazard,

$$\lambda_\bullet(t;\boldsymbol{z}) = \boldsymbol{z}\,\boldsymbol{\beta}_\bullet(t).$$

Let $\hat{\beta}_k(t)\ \forall k$ be the estimator of the regression function of the *k*th cause-specific hazard. A natural estimator of the regression function of the all-cause hazard would the sum of the estimators of the regression functions of all cause-specific hazards. The covariate adjusted cumulative hazard due to the *k*th cause is given by $\hat{A}_k(t;\boldsymbol{z}) = \int_0^t \boldsymbol{z}\,\hat{\boldsymbol{\beta}}_k(t)\,du$. The overall survival can be estimated by $\hat{S}(t;\boldsymbol{z}) = \exp(-\sum_k \hat{A}_k(t;\boldsymbol{z}))$.

In this thesis, we apply the direct adjustment method to give inference on assessing treatment effect with respect to a specific type of failure. Suppose that *K* types of failure exist. In order to simply the notations, we consider two treatments only, but the results can be easily

extended to a setting with more than two treatments. Let *X* indicates the treatment group index, where $X \in \{1, 2\}$. We define

$$\boldsymbol{z}_l^i = (1, X = i, \widetilde{\boldsymbol{Z}}_l), \qquad i = 1, 2; l = 1, \cdots, n. \tag{2.17}$$

Let $A_{ik}(t)$ be the direct adjusted cumulative hazard of the *i*th treatment for the *k*th type of failure. It can be estimated by

$$\hat{A}_{ik}(t) = \frac{1}{n}\sum_{l=1}^{n} \hat{A}_k(t; \boldsymbol{z}_l^i). \tag{2.18}$$

Let $\gamma_k(t)$ be the difference in the adjusted cumulative hazard between two treatments for the *k*th type of failure,

$$\hat{\gamma}_k(t) = \hat{A}_{1k}(t) - \hat{A}_{2k}(t). \tag{2.19}$$

Variances of $\hat{A}_{ik}(t)$ and $\hat{\gamma}(t)$ can be estimated by

$$\hat{\text{var}}[\hat{A}_{ik}(t)] = \left(\frac{1}{n}\sum_{l=1}^{n} \boldsymbol{z}_l^i\right)\hat{\boldsymbol{\Sigma}}_k(t)\left(\frac{1}{n}\sum_{l=1}^{n} \boldsymbol{z}_l^i\right)^T \tag{2.20}$$

$$\hat{\text{var}}[\hat{\gamma}_k(t)] = \left(\frac{1}{n}\sum_{l=1}^{n} \boldsymbol{z}_l^1 - \frac{1}{n}\sum_{l=1}^{n} \boldsymbol{z}_l^2\right)\hat{\boldsymbol{\Sigma}}_k(t)\left(\frac{1}{n}\sum_{l=1}^{n} \boldsymbol{z}_l^1 - \frac{1}{n}\sum_{l=1}^{n} \boldsymbol{z}_l^2\right)^T \tag{2.21}$$

We define $F_{ik}(t)$ to be the cumulative incidence function of treatment *i* for the *k*th type of failure. The direct adjusted estimator of $F_{ik}(t)$ is

$$\hat{F}_{ik}(t) = n^{-1}\sum_{l=1}^{n} \hat{F}_k(t; \boldsymbol{z}_l^i) = n^{-1}\sum_{l=1}^{n}\int_0^t \hat{S}(u-; \boldsymbol{z}_l^i) d\hat{A}_k(u; \boldsymbol{z}_l^i) \tag{2.22}$$

Its variance can be estimated by

$$\hat{\text{var}}[\hat{F}_{ik}(t)] = \int_0^t J(u)\hat{\boldsymbol{v}}(u)^T \boldsymbol{Y}^-(u) diag\boldsymbol{N}_k(u)(\boldsymbol{Y}^-(u))^T \hat{\boldsymbol{v}}(u)du + \sum_{l\neq k}\int_0^t J(u)\hat{\boldsymbol{w}}(u)^T \boldsymbol{Y}^-(u) diag\boldsymbol{N}_l(u)(\boldsymbol{Y}^-(u))^T \hat{\boldsymbol{w}}(u)du \quad (2.23)$$

where

$$\hat{\boldsymbol{v}}(u) = n^{-1}\sum_{l=1}^{n}(\hat{S}(u-;\boldsymbol{z}) - \hat{F}_k(t;\boldsymbol{z}_l^i) + \hat{F}_k(u;\boldsymbol{z}_l^i))\boldsymbol{z}_l^i \quad (2.24)$$

$$\hat{\boldsymbol{w}}(u) = n^{-1}\sum_{l=1}^{n}(\hat{F}_k(t;\boldsymbol{z}_l^i) - \hat{F}_k(u;\boldsymbol{z}_l^i))\boldsymbol{z}_l^i \quad (2.25)$$

The difference of cumulative incidence functions between two treatments is often of study interest. Let $\mu_k(t) = F_{1k}(t) - F_{2k}(t)$. Obviously, $\hat{\mu}_k(t) = \hat{F}_{1k}(t) - \hat{F}_{2k}(t)$. We have derived the variance formula of $\hat{\mu}_k(t)$. The estimated variance is given by

$$\hat{\text{var}}[\hat{\mu}_k(t)] = \int_0^t J(u)\hat{\boldsymbol{\eta}}(u)^T \boldsymbol{Y}^-(u) diag\boldsymbol{N}_k(u)(\boldsymbol{Y}^-(u))^T \hat{\boldsymbol{\eta}}(u)du + \sum_{l\neq k}\int_0^t J(u)\hat{\boldsymbol{\rho}}(u)^T \boldsymbol{Y}^-(u) diag\boldsymbol{N}_l(u)(\boldsymbol{Y}^-(u))^T \hat{\boldsymbol{\rho}}(u)du \quad (2.26)$$

where

$$\hat{\boldsymbol{\eta}}(u) = n^{-1}\sum_{l=1}^{n}(\hat{S}(u-;\boldsymbol{z}_l^1) - \hat{F}_k(t;\boldsymbol{z}_l^1) + \hat{F}_k(u;\boldsymbol{z}_l^1))\boldsymbol{z}_l^1 - n^{-1}\sum_{l=1}^{n}(\hat{S}(u-;\boldsymbol{z}_l^2) - \hat{F}_k(t;\boldsymbol{z}_l^2) + \hat{F}_k(u;\boldsymbol{z}_l^2))\boldsymbol{z}_l^2 \quad (2.27)$$

$$\hat{\boldsymbol{\rho}}(u) = n^{-1}\sum_{l=1}^{n}(\hat{F}_k(u;\boldsymbol{z}_l^1) - \hat{F}_k(t;\boldsymbol{z}_l^1))\boldsymbol{z}_l^1 - n^{-1}\sum_{l=1}^{n}(\hat{F}_k(t;\boldsymbol{z}_l^2) + \hat{F}_k(u;\boldsymbol{z}_l^2))\boldsymbol{z}_l^2 \quad (2.28)$$

CHAPTER III. SIMULATION STUDY

The settings considered in the simulation study include two treatments (1 and 2). Both discrete and continuous covariates have been simulated in the study. Treatment assignment was determined by generating a Bernoulli random variable with the probability 0.5 to be assigned to treatment 1, and the probability 0.5 to be assigned to treatment 2.

First, we conducted simulation on settings with discrete covariates. We let treatment 1 be the better treatment option with the baseline cumulative hazard $0.1t$, an term $0.02t^2$ should be added to the cumulative hazard if the observation is in treatment 2 group. Three covariates, Z_1, Z_2 and Z_3 were generated from Bernoulli distributions with the probability 0.5 to take the value 1. Given t, the cumulative hazard will be increased by $0.05t$, $0.1t^2$, $0.3t^2$ if Z_1, Z_2 or Z_3 takes the value 1, respectively.

In summary, given $\boldsymbol{z}$, the cumulative hazard with treatment 1 is

$$A_1(t;\mathrm{z}) = 0.1t + 0.05tZ_1 + 0.1t^2Z_2 + 0.3t^2Z_3, \qquad 3.1$$

and the cumulative hazard with treatment 2 is

$$A_2(t;z) = 0.1t + 0.02t^2 + 0.05tZ_1 + 0.1t^2Z_2 + 0.3t^2Z_3. \qquad 3.2$$

Utilizing the relation $S(t;z) = \exp\{-A(t;z)\}$, we applied the inverse transform method to generate failure time T. First, a random unit, U, needs to be generated. Then the failure time T, given treatment 1, shall be obtained by the following formulas with various values of $\boldsymbol{z}$,

i. $Z_1(t) = 0, Z_2(t) = 0, Z_3(t) = 0$, $\quad T = -10\log(U)$,

ii. $Z_1(t) = 0, Z_2(t) = 0, Z_3(t) = 1$, $\quad T = \dfrac{-0.1 + \sqrt{0.01 - 1.2\log(U)}}{0.6}$,

iii. $Z_1(t)=0, Z_2(t)=1, Z_3(t)=0$, $T=\dfrac{-0.1+\sqrt{0.01-0.4\log(U)}}{0.2}$,

iv. $Z_1(t)=0, Z_2(t)=1, Z_3(t)=1$, $T=\dfrac{-0.1+\sqrt{0.01-1.6\log(U)}}{0.8}$,

v. $Z_1(t)=1, Z_2(t)=0, Z_3(t)=0$, $T=\dfrac{-0.1+\sqrt{0.01-0.08\log(U)}}{0.04}$,

vi. $Z_1(t)=1, Z_2(t)=0, Z_3(t)=1$, $T=\dfrac{-0.1+\sqrt{0.01-1.28\log(U)}}{0.64}$,

vii. $Z_1(t)=1, Z_2(t)=1, Z_3(t)=0$, $T=\dfrac{-0.1+\sqrt{0.01-0.48\log(U)}}{0.24}$,

viii. $Z_1(t)=1, Z_2(t)=1, Z_3(t)=1$, $T=\dfrac{-0.1+\sqrt{0.01-1.68\log(U)}}{0.84}$.

Given treatment 2, the failure time can be generated by the following formulas,

i. $Z_1(t)=0, Z_2(t)=0, Z_3(t)=0$, $T=\dfrac{-0.1+\sqrt{0.01-0.08\log(U)}}{0.04}$,

ii. $Z_1(t)=0, Z_2(t)=0, Z_3(t)=1$, $T=\dfrac{-0.1+\sqrt{0.01-1.28\log(U)}}{0.64}$,

iii. $Z_1(t)=0, Z_2(t)=1, Z_3(t)=0$, $T=\dfrac{-0.1+\sqrt{0.01-0.48\log(U)}}{0.24}$,

iv. $Z_1(t)=0, Z_2(t)=1, Z_3(t)=1$, $T=\dfrac{-0.1+\sqrt{0.01-1.68\log(U)}}{0.84}$,

v. $Z_1(t)=1, Z_2(t)=0, Z_3(t)=0$, $T=\dfrac{-0.15+\sqrt{0.0225-0.08\log(U)}}{0.04}$,

vi. $Z_1(t)=1, Z_2(t)=0, Z_3(t)=1$, $\quad T=\dfrac{-0.15+\sqrt{0.0225-1.28\log(U)}}{0.64}$,

vii. $Z_1(t)=1, Z_2(t)=1, Z_3(t)=0$, $\quad T=\dfrac{-0.15+\sqrt{0.0225-0.48\log(U)}}{0.24}$,

viii. $Z_1(t)=1, Z_2(t)=1, Z_3(t)=1$, $\quad T=\dfrac{-0.15+\sqrt{0.0225-1.68\log(U)}}{0.84}$.

The censoring time was generated from the Uniform distribution [0, *a*]. The value of *a* was chosen to yield the censoring rates 25% and 50%. We considered sample size 100 and 200, 1000 replicates were simulated for each setting. The simulation results are shown in Table 1 and Table 2. In the tables, we report the difference between the average of the estimated directed survival probabilities and the true values, the empirical standard errors, the averages of the estimated standard errors, and the 95% coverage. For 95% coverage, we computed three types of confidence intervals. They are linear, log-log transformed and arcsine-square-root transformed confidence intervals.

For the second set of simulation, we considered continuous covariates. We still let treatment 1 be the better treatment option with the baseline cumulative hazard t^2, the cumulative hazard for treatment 2 was set to be $0.4t^2$ higher. The covariates Z_1, Z_2 and Z_3 were generated from a standard Normal distribution. We further set, given ***z***, the cumulative hazard with treatment 1 is

$$A_1(t;z)=t^2+0.2tZ_1+0.3t^2Z_2+0.1tZ_3, \qquad 3.3$$

and the cumulative hazard with treatment 2 is

$$A_2(t;z)=t^2+0.4t^2+0.2tZ_1+0.3t^2Z_2+0.1tZ_3. \qquad 3.4$$

The failure time was generated using the inverse transform method. First, generate a random unit, *U*. The failure time *T*, given treatment 1, is obtained by

$$T = \frac{-(0.2Z_1 + 0.1Z_3) + \sqrt{(0.2Z_1 + 0.1Z_3)^2 - 4(1 + 0.3Z_2)\log(U)}}{2(1 + 0.3Z_2)}.$$

Given treatment 2, the failure time can be generated by

$$T = \frac{-(0.2Z_1 + 0.1Z_3) + \sqrt{(0.2Z_1 + 0.1Z_3)^2 - 4(1.4 + 0.3Z_2)\log(U)}}{2(1.4 + 0.3Z_2)}.$$

The censoring time was generated from the Uniform distribution [0, *a*]. We simulated settings with the censoring rates 25%, 50%, and the sample size 100, 200. 1000 replicates were utilized. The simulation results are shown in Table 3 and Table 4.

According to Table 1-4, the average estimated standard errors are quite close to the empirical standard errors when the regression model contains only continuous covariates, but they are slightly smaller than the empirical standard errors for models with discrete covariates. Through all settings, the log-log transformed confidence interval consistently performs better than the linear or arcsine-square-root transformed confidence interval, and should be strongly recommended.

Table 1: Simulation results for discrete covariates with censoring rate 25%

t	0.75	1	1.25	1.5	1.75	2.0
n = 100						
$S_1(t)$						
Bias	0.005	0.008	0.013	0.016	0.017	0.021
Empirical SE	0.056	0.065	0.071	0.075	0.077	0.074
Estimated SE	0.053	0.061	0.067	0.071	0.072	0.072
Linear-Cov.	0.901	0.920	0.918	0.908	0.919	0.927
LogLog-Cov.	0.962	0.955	0.951	0.936	0.936	0.944
Arcsin-Cov.	0.929	0.937	0.930	0.917	0.926	0.931
$S_2(t)$						
Bias	0.006	0.008	0.010	0.012	0.012	0.017
Empirical SE	0.056	0.065	0.073	0.074	0.077	0.077
Estimated SE	0.055	0.064	0.069	0.072	0.072	0.072
Linear-Cov.	0.928	0.925	0.916	0.929	0.922	0.921
LogLog-Cov.	0.963	0.954	0.947	0.949	0.939	0.938
Arcsin-Cov.	0.943	0.941	0.925	0.936	0.926	0.925
n = 200						
$S_1(t)$						
Bias	0.003	0.005	0.007	0.008	0.008	0.010
Empirical SE	0.039	0.045	0.050	0.052	0.052	0.052
Estimated SE	0.038	0.044	0.048	0.050	0.051	0.051
Linear-Cov.	0.926	0.938	0.932	0.931	0.935	0.934
LogLog-Cov.	0.959	0.949	0.947	0.951	0.948	0.945
Arcsin-Cov.	0.948	0.948	0.936	0.939	0.936	0.937
$S_2(t)$						
Bias	0.004	0.004	0.006	0.007	0.007	0.006
Empirical SE	0.041	0.048	0.050	0.053	0.054	0.054
Estimated SE	0.039	0.045	0.049	0.051	0.051	0.050
Linear-Cov.	0.920	0.937	0.947	0.939	0.929	0.928
LogLog-Cov.	0.956	0.954	0.965	0.955	0.937	0.934
Arcsin-Cov.	0.934	0.948	0.954	0.942	0.935	0.930

Table 2: Simulation results for discrete covariates with censoring rate 50%

t	0.75	1	1.25	1.5	1.75	2.0
n = 100						
$S_1(t)$						
Bias	0.006	0.010	0.015	0.020	0.023	0.030
Empirical SE	0.058	0.068	0.077	0.083	0.088	0.091
Estimated SE	0.055	0.065	0.073	0.078	0.083	0.086
Linear-Cov.	0.908	0.925	0.917	0.913	0.922	0.919
LogLog-Cov.	0.958	0.954	0.949	0.941	0.951	0.940
Arcsin-Cov.	0.928	0.938	0.931	0.921	0.934	0.925
$S_2(t)$						
Bias	0.006	0.009	0.012	0.016	0.019	0.025
Empirical SE	0.058	0.070	0.079	0.081	0.086	0.092
Estimated SE	0.057	0.067	0.074	0.079	0.082	0.085
Linear-Cov.	0.925	0.919	0.915	0.931	0.929	0.923
LogLog-Cov.	0.962	0.948	0.948	0.941	0.948	0.939
Arcsin-Cov.	0.948	0.941	0.932	0.953	0.934	0.925
n = 200						
$S_1(t)$						
Bias	0.003	0.006	0.007	0.008	0.011	0.014
Empirical SE	0.041	0.048	0.054	0.058	0.059	0.062
Estimated SE	0.039	0.047	0.052	0.056	0.058	0.059
Linear-Cov.	0.931	0.929	0.926	0.927	0.934	0.932
LogLog-Cov.	0.956	0.952	0.943	0.940	0.947	0.947
Arcsin-Cov.	0.946	0.935	0.931	0.933	0.936	0.936
$S_2(t)$						
Bias	0.005	0.005	0.007	0.008	0.009	0.009
Empirical SE	0.043	0.050	0.054	0.060	0.061	0.062
Estimated SE	0.041	0.048	0.053	0.056	0.058	0.059
Linear-Cov.	0.928	0.938	0.945	0.931	0.932	0.922
LogLog-Cov.	0.945	0.950	0.956	0.939	0.945	0.928
Arcsin-Cov.	0.938	0.948	0.952	0.937	0.938	0.925

Table 3: Simulation results for continuous covariates with censoring rate 25%

t	0.2	0.4	0.6	0.8	1.0
n = 100					
$S_1(t)$					
Bias	-0.008	0.001	0.011	0.018	0.024
Empirical SE	0.067	0.051	0.066	0.077	0.080
Estimated SE	0.026	0.050	0.067	0.077	0.081
Linear-Coverage	0.885	0.930	0.935	0.931	0.939
LogLog-Coverage	0.848	0.960	0.959	0.946	0.956
Arcsin-Coverage	0.922	0.948	0.949	0.942	0.942
$S_2(t)$					
Bias	-0.008	0.003	0.010	0.016	0.022
Empirical SE	0.068	0.056	0.071	0.078	0.075
Estimated SE	0.031	0.057	0.072	0.076	0.073
Linear-Coverage	0.919	0.938	0.944	0.936	0.921
LogLog-Coverage	0.893	0.970	0.971	0.949	0.932
Arcsin-Coverage	0.937	0.956	0.956	0.939	0.928
n = 200					
$S_1(t)$					
Bias	-0.006	-0.003	0.000	0.004	0.009
Empirical SE	0.021	0.037	0.049	0.054	0.056
Estimated SE	0.020	0.036	0.048	0.054	0.055
Linear-Coverage	0.918	0.936	0.935	0.944	0.945
LogLog-Coverage	0.920	0.944	0.947	0.951	0.956
Arcsin-Coverage	0.952	0.945	0.944	0.947	0.946
$S_2(t)$					
Bias	-0.002	0.001	0.003	0.008	0.008
Empirical SE	0.023	0.041	0.052	0.055	0.053
Estimated SE	0.022	0.041	0.051	0.054	0.050
Linear-Coverage	0.916	0.948	0.934	0.945	0.933
LogLog-Coverage	0.958	0.955	0.941	0.953	0.941
Arcsin-Coverage	0.938	0.951	0.938	0.948	0.933

Table 4: Simulation results for continuous covariates with censoring rate 50%

t	0.2	0.4	0.6	0.8	1.0
n = 100					
$S_1(t)$					
Bias	-0.008	0.001	0.013	0.024	0.047
Empirical SE	0.067	0.053	0.072	0.089	0.113
Estimated SE	0.027	0.053	0.073	0.088	0.113
Linear-Coverage	0.880	0.932	0.926	0.923	0.920
LogLog-Coverage	0.840	0.962	0.969	0.944	0.948
Arcsin-Coverage	0.919	0.946	0.946	0.935	0.926
$S_2(t)$					
Bias	-0.008	0.003	0.013	0.020	0.037
Empirical SE	0.068	0.059	0.078	0.090	0.101
Estimated SE	0.032	0.059	0.078	0.087	0.097
Linear-Coverage	0.911	0.925	0.943	0.928	0.933
LogLog-Coverage	0.891	0.967	0.962	0.950	0.946
Arcsin-Coverage	0.935	0.951	0.952	0.933	0.939
n = 200					
$S_1(t)$					
Bias	-0.005	-0.002	0.002	0.007	0.016
Empirical SE	0.022	0.038	0.052	0.061	0.067
Estimated SE	0.020	0.038	0.052	0.061	0.068
Linear-Coverage	0.912	0.941	0.947	0.941	0.948
LogLog-Coverage	0.918	0.953	0.962	0.959	0.960
Arcsin-Coverage	0.950	0.949	0.956	0.948	0.954
$S_2(t)$					
Bias	-0.002	0.001	0.005	0.011	0.016
Empirical SE	0.023	0.043	0.057	0.062	0.064
Estimated SE	0.023	0.043	0.055	0.061	0.062
Linear-Coverage	0.913	0.936	0.923	0.935	0.931
LogLog-Coverage	0.964	0.956	0.939	0.946	0.937
Arcsin-Coverage	0.948	0.946	0.928	0.944	0.930

CHAPTER IV. EXAMPLES

4.1 Example 1

The goal of study is to compare three types of stem cell transplantation in treating follicular lymphoma patients. The study cohort consists of 904 follicular lymphoma patients who received either allogeneic or autologous transplant between 1990 and 1999 and reported to the International Bone Marrow Transplant Registry. Among these 904 patients, 176 received allogeneic transplant, 131 received purged autologous transplant, and 597 received unpurged autologous transplants. The median follow up times for survivors are 36, 49 and 41 months for allogeneic, purged autologous and unpurged autologous transplant recipients. Distribution of risk factors in three transplant groups is shown in Table 5. It seems that that the purged autologous transplant group contained more healthy patients than the other two groups. In the purge autologous group, relatively more patients had high Karnofsky scores (87%) and were in the early stage of disease (53%). Differences in patients' baseline profiles between transplant groups need to be adjusted when one compares three types of transplantation. This data set was originally analyzed by Besien et al. (2003). Because the hazards between transplant groups are apparently nonproportional, Besien et al. considered a Cox model stratified on transplant groups. Aalen's model gives more flexibility on modeling covariate effects. We wish to reanalyze the data set and we are interested in comparison in the survival between transplant groups at 3-year and 5-year after transplant.

The macro described in Section 2.2 was applied to give us the comparison results. First, we made a SAS data set that included the following variables: "time"=failure time; "death"=death indicator; "transplant"=1 for unpurged auto transplant, 2 for purged auto transplant, 3 for allogeneric transplant; "stage"=1 if disease is in advanced stage, 0 otherwise;

"chemo1"=1 if chemosensitivity is resistant; "chemo2"=1 if chemosensitivity is untreated / unknown, 0 otherwise; "LHD1"=1 if LHD is abnormal; "LHD2"=1 if LHD is unknown; "kscore"=1 if Karnofsky is 90%- 100%; "DX2T1"=1 if interval from diagnosis is 1-2 years; "DX2T2"=1 if interval from diagnosis is more than 2 years; and "age"=1 if older than 40; "year1"=1 if year of diagnosis between 1994-1996 ; "year2"=1 if year of diagnosis greater or equal to 1997. The macro was loaded by **%include** statement and invoked by

%estimate (data, time, event, group, stage chemo1 chemo2 LHD1 LHD2 kscore DX2T1 DX2T2 age year1 year2 , out).

The direct adjusted survival curves of three types of transplant have been plotted in Figure 1. In the figure, two autologous transplants yield better survival outcome than allogeneic transplant, during the majority study period. Among two autologous transplants, they have similar survival outcome within 1 year post transplant, while purged transplant give higher survival rate after 1 year. Figure 2-4 shows the differences between any two transplants and their confidence intervals. These figures suggest obvious changes in treatment effect over time.

We further select the time points, 3 year and 5 year, to assess effects of transplants. Comparison of survival rates between any two transplants, at the given time points, is given in Table 6. At both time points, unpurged autologous transplant yields lower survival probabilities than purged autologous transplant. However, the effect is only marginally significant (*p*-values: 0.1075 and 0.0869, respectively). Compared to allogeneic transplant, two autologous transplants have significant better result in terms of three-year survival rate (*p*-values: 0.0087 and 0.0019, respectively). Two autologous transplant still outperform allogeneic transplant in terms of five year survival rate (*p*-values: 0.0694 and 0.0465, respectively), however, the magnitude has been reduced.

Table 5: Groups and risk factors for transplantation of follicular lymphoma patients

	Unpurged		Purged		Allogeneic	
Total	597		131		176	
Stage						
Early	350	(59%)	69	(53%)	85	(48%)
Advanced	247	(41%)	62	(47%)	91	(52%)
Chemosensitivity						
Sensitivity	488	(82%)	111	(85%)	118	(67%)
Resistant	66	(11%)	14	(11%)	31	(18%)
Untreated/unknown	43	(7%)	6	(4%)	27	(15%)
LHD						
Normal	389	(65%)	55	(42%)	117	(67%)
Abnormal	174	(29%)	27	(21%)	44	(25%)
Unknown	34	(6%)	49	(37%)	15	(8%)
Karnofsky score						
No more than 80%	183	(31%)	17	(13%)	60	(34%)
90%-100%	414	(69%)	114	(87%)	116	(66%)
Interval from diagnosis						
Less than 1 yr	115	(19%)	27	(21%)	26	(15%)
1-2 yrs	156	(16%)	32	(24%)	56	(32%)
More than 2 yrs	326	(55%)	72	(55%)	94	(53%)
Age group						
No older than 40	100	(17%)	22	(17%)	77	(44%)
Older than 40	497	(83%)	109	(83%)	99	(56%)

Table 6: Direct adjusted survivals of unpurged-auto (S1), purged-auto (S2) and allogenic (S3)

	S1	SE	S2	SE	S3	SE
36 months	0.651	0.020	0.708	0.041	0.546	0.038
60 months	0.550	0.025	0.626	0.049	0.517	0.042

	S1-S2	SE(S1-S2)	*p*-value
36 months	-0.057	0.046	0.1075
60 months	-0.075	0.055	0.0869

	S1-S3	SE(S1-S3)	*p*-value
36 months	0.105	0.044	0.0087
60 months	0.033	0.049	0.0694

	S2-S3	SE(S2-S3)	*p*-value
36 months	0.162	0.056	0.0019
60 months	0.108	0.064	0.0465

Figure 1: Direct adjusted survivals of unpurged-autologous, purged-autologous and allogenic transplants

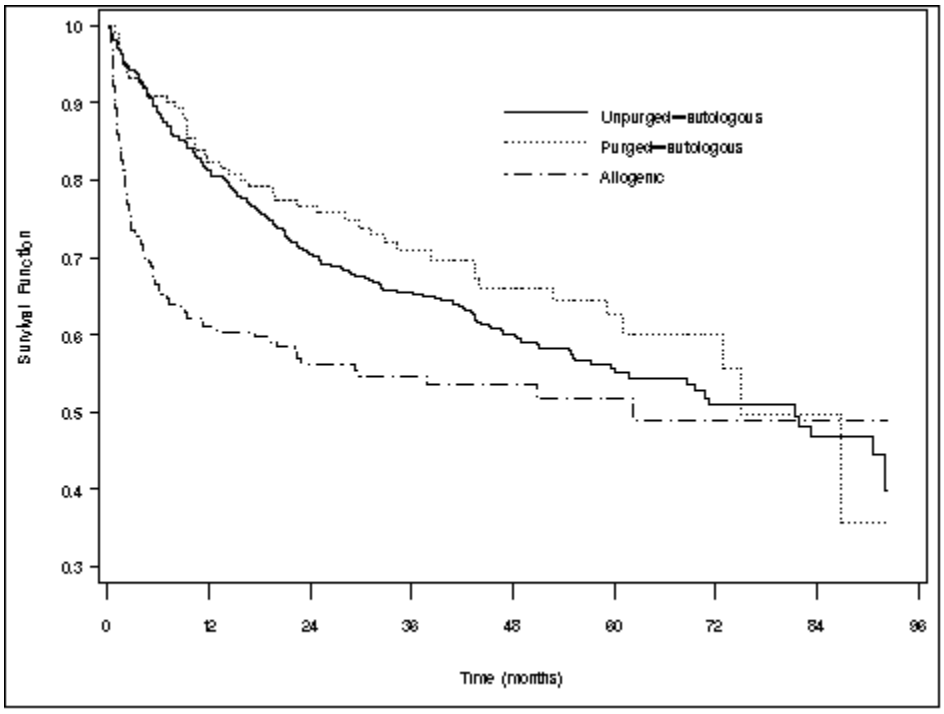


Figure 2: Difference of direct adjusted survivals of unpurged-autologous and purged-autologous transplants

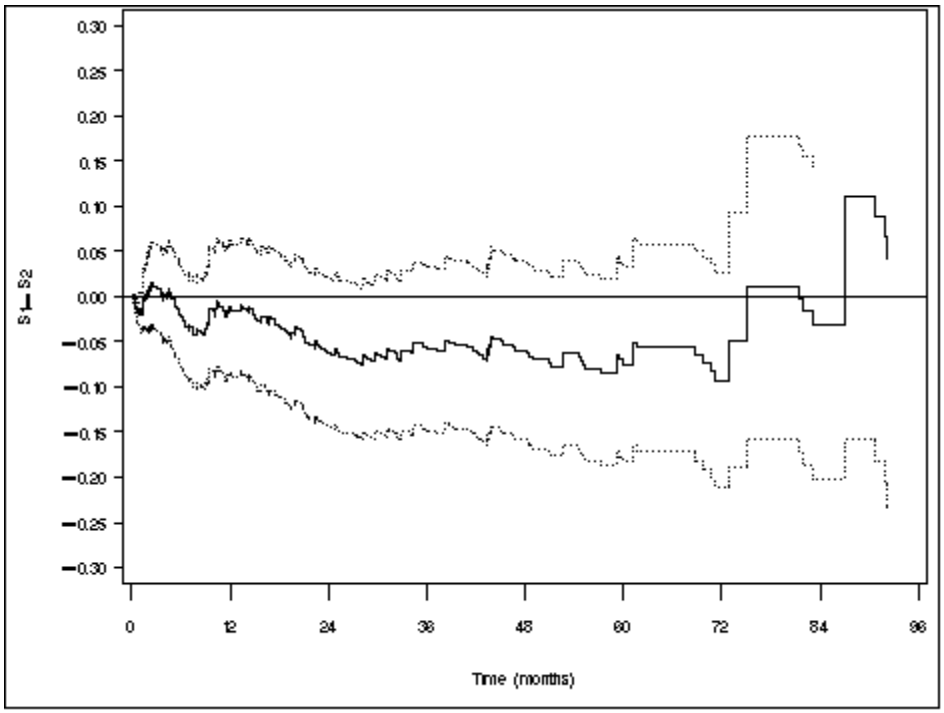


Figure 3: Difference of direct adjusted survivals of unpurged-autologous and allogenic transplants

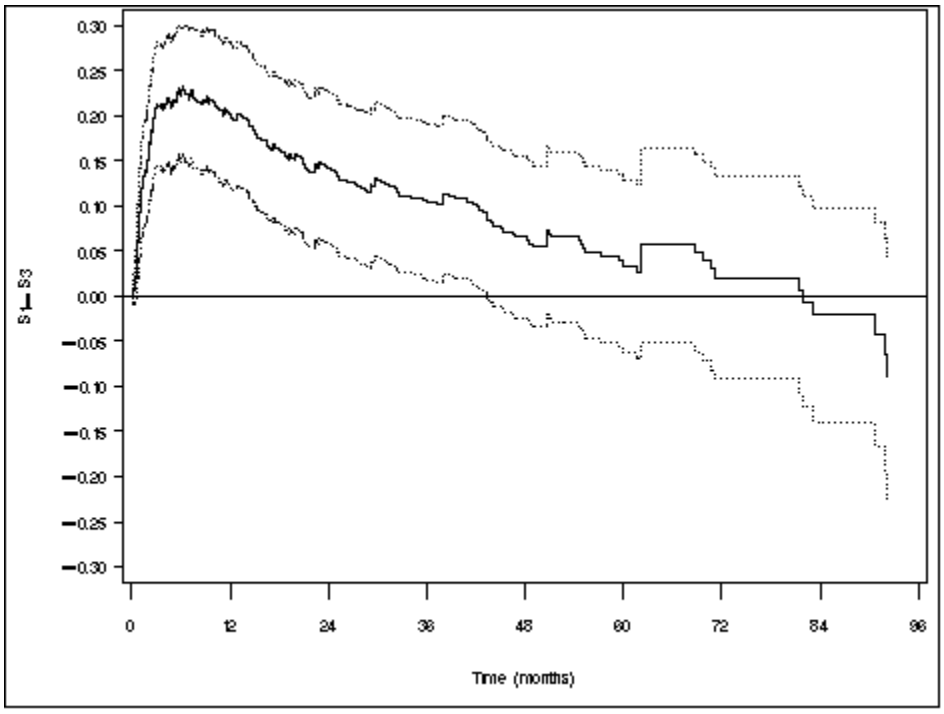
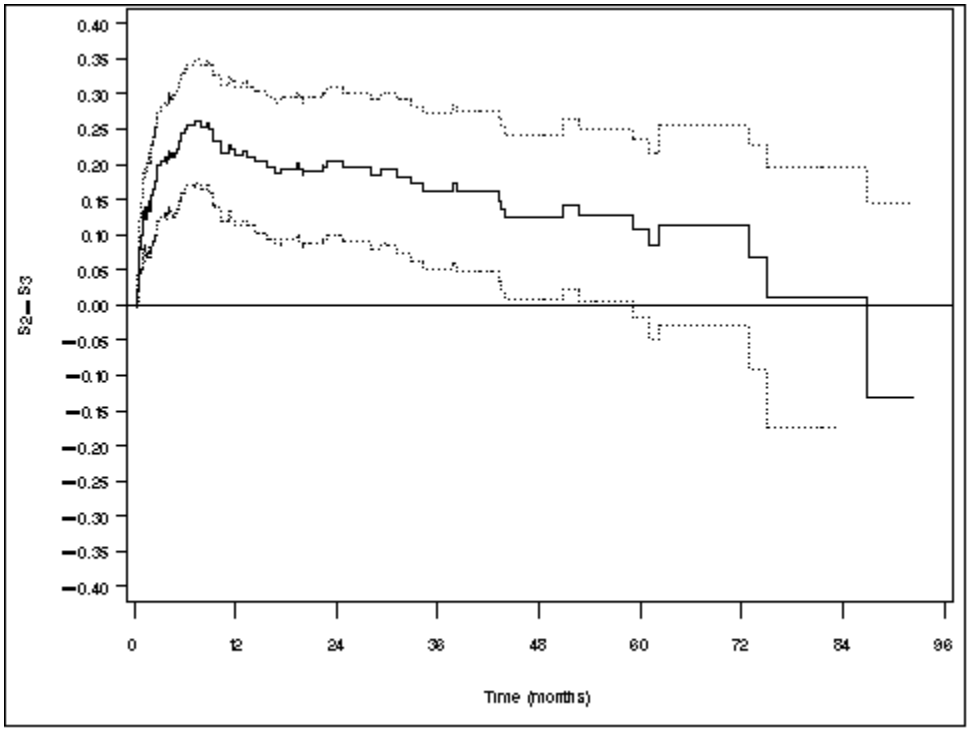


Figure 4: Difference of direct adjusted survivals of purged-autologous and allogenic transplants



4.2 Example 2

In this study, we would compare the direct adjusted cumulative incidence probabilities between two treatment options (mastectomy vs. breast conserving surgery) for early staged breast cancer patients. The study cohort was selected from the published SEER cancer case registry. The Surveillance, Epidemiology, and End Results (SEER) program collects all cancer cases occurring in the participating SEER sites and periodically publish the database to give source to clinicians and epidemiologist who conduct various cancer-related research projects. The SEER registry was initiated in 1973 with 9 sites and has now been expanded to 13 sites, covering 26% of the US population. In our study, female breast cancer cases were selected if a patient was diagnosed with stage I/II breast cancer between 1991 and 1996, resided in Atlanta at diagnosis, was 30-79 years of age at time of diagnosis, and her breast cancer was the first malignant primary cancer and was microscopically confirmed, and either mastectomy or breast conserving surgery (BCS) was conducted. A total of 3760 patients entered the study cohort. The cut-off date for the database used in this study is December 31, 2002. Most of the patients were followed till this date, but a small percentage of patients were lost to follow-up before this date. The median follow up time is 92 months. During the follow-up, 929 deaths were recorded: 598 died from breast cancer, 162 died from vascular disease and 169 died from other causes.

In this example, three competing risks exist: breast cancer, vascular disease and other cause. Mastectomy and BCS are the surgical procedures that patients received. Mastectomy is the surgical procedure of removing the affected breast. BCS (or lumpectomy) is to remove only the lump (tumor), which is considered to be non-invasive compared to mastectomy. Prognostic factors that we identified from SEER database include demographical characteristics, patient's age at diagnosis and race, and tumor-related characteristics, tumor size, lymph node status, tumor

grade and extension. Table 7 shows the distribution of the risk factors in two treatment groups. It is obvious that the distribution of tumor-related factors is very different between two groups, the cohort receiving mastectomy have much worse profiles than the cohort receiving BCS. It should be important to address the problem of the imbalance in distribution of tumor-related factors when comparing these two surgical procedures.

When fitting Aalen's models on all cause-specific hazards, we include the following factors: race (White, Black and Other), age at diagnosis (30-49, 50-64 and 65-79), tumor size (< 2 cm, 2-5 cm and > 5 cm), lymph node status (negative, positive and unknown/unexamined), tumor grade (1, 2 and 3 / 4), tumor extension (confined and invasive).

The results are given for the predetermined time points, 5 years and 8 years. Figures 5, 7 and 9 give the direct adjusted cumulative incidence curves for both types of surgery on cancer, vascular disease, and other cause, respectively. Differences between two direct adjusted cumulative incidence curves and their 95% confidence intervals, for all competing risks, are given in Figures 6, 8 and 10. Figure 5 and 6 show that mastectomy and BCS have similar outcome on cancer. Figure 7-10 show that mastectomy has higher cumulative incidence than BCS for vascular disease and other cause, but the effect is only marginally significant at the selected times points. Direct adjusted cumulative incidence estimates and their estimated standard errors, as well as the results on treatment comparison at the selected times points are given in Table 8. Finally, we give the results on direct adjusted cumulative cause-specific hazards in Figures 11-14.. It is interesting that treatment effect is significant with respect to direct adjusted cause-specific hazard, but insignificant or marginal significant with respect to direct adjusted cumulative incidences. Because of the drawback of cause-specific hazard in the

setting of competing risks, the results on the direct adjusted cumulative incidences are recommended.

Table 7: Causes, treatments and risk factors for breast cancer patients

	Mastectomy		BCS	
Total	2020	(53.7%)	1740	(46.3%)
Cause				
Censored	1422	(70.4%)	1409	(81%)
Cancer	386	(19.1%)	212	(12.2%)
Vascular Disease	107	(5.3%)	55	(3.1%)
Other	105	(5.2%)	64	(3.7%)
Age groups				
30-49	717	(35.5%)	650	(37.3%)
50-64	691	(34.2%)	673	(38.7%)
65-79	612	(30.3%)	417	(24.0%)
Race				
White	1480	(73.3%)	1315	(75.6%)
Black	485	(24%)	383	(22%)
Other	55	(2.7%)	42	(2.4%)
Size group				
<2 cm	917	(45.4%)	1190	(68.4%)
2-5 cm	1070	(53.0%)	546	(31.4%)
>5 cm	33	(1.6%)	4	(0.2%)
Node				
Negative	1194	(59.1%)	1208	(69.4%)
Positive	783	(38.7%)	417	(23.9%)
Unknown/Unexamined	43	(2.1%)	115	(6.6%)
Grade				
1	220	(10.9%)	346	(19.9%)
2	839	(41.5%)	722	(41.5%)
3 /4	961	(47.6%)	672	(38.6%)
Ext				
Confined	1932	(95.6%)	1721	(98.9%)
Invasive	88	(4.4%)	19	(1.1%)
Diagnosis Year				
1991	331	(16.4%)	177	(10.2%)
1992	344	(17.0%)	207	(11.9%)
1993	336	(16.6%)	276	(15.9%)
1994	327	(16.2%)	302	(17.3%)
1995	352	(17.4%)	353	(20.3%)
1996	330	(16.3%)	425	(24.4%)

Table 8: Direct adjusted cumulative incidence functions for breast cancer patients

Cause	CIF1	SE	CIF2	SE
Cancer				
60 months	0.102	0.006	0.101	0.007
96 months	0.153	0.008	0.146	0.008
Vascular Disease				
60 months	0.021	0.003	0.015	0.002
96 months	0.043	0.004	0.032	0.004
Other				
60 months	0.028	0.003	0.019	0.003
96 months	0.047	0.004	0.036	0.004

Cause	(CIF1-CIF2)	SE(CIF1-CIF2)	*p*-value
Cancer			
60 months	0.001	0.009	0.456
96 months	0.007	0.012	0.281
Vascular Disease			
60 months	0.006	0.004	0.067
96 months	0.011	0.007	0.058
Other			
60 months	0.009	0.005	0.036
96 months	0.011	0.007	0.058

Figure 5 : Direct adjusted cumulative incidence functions of mastectomy and BCS on cancer

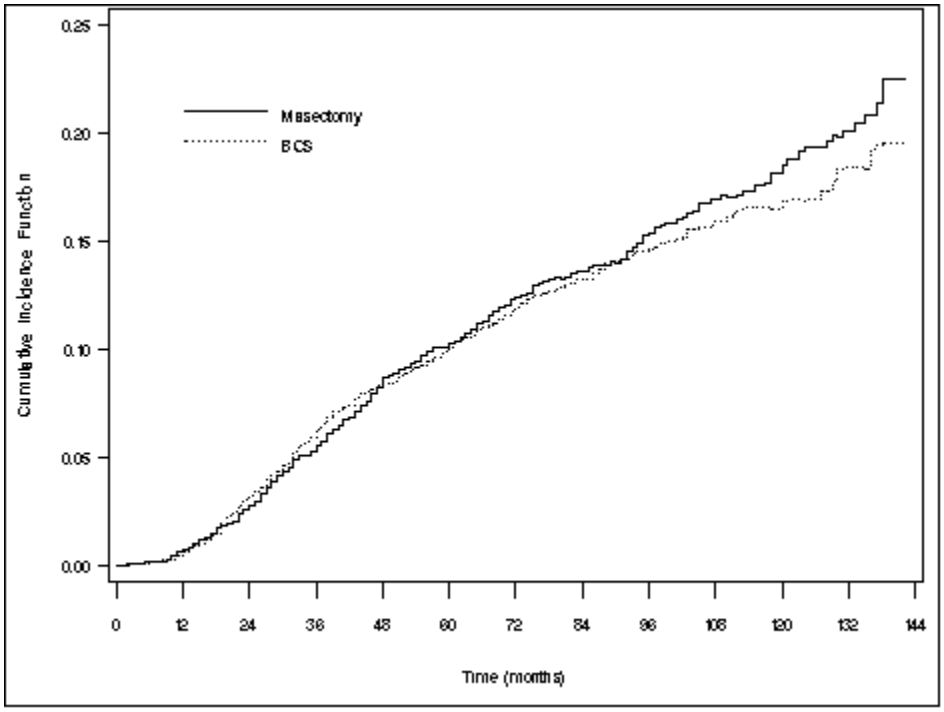


Figure 6 : Difference between direct adjusted cumulative incidence functions of mastectomy and BCS on Cancer

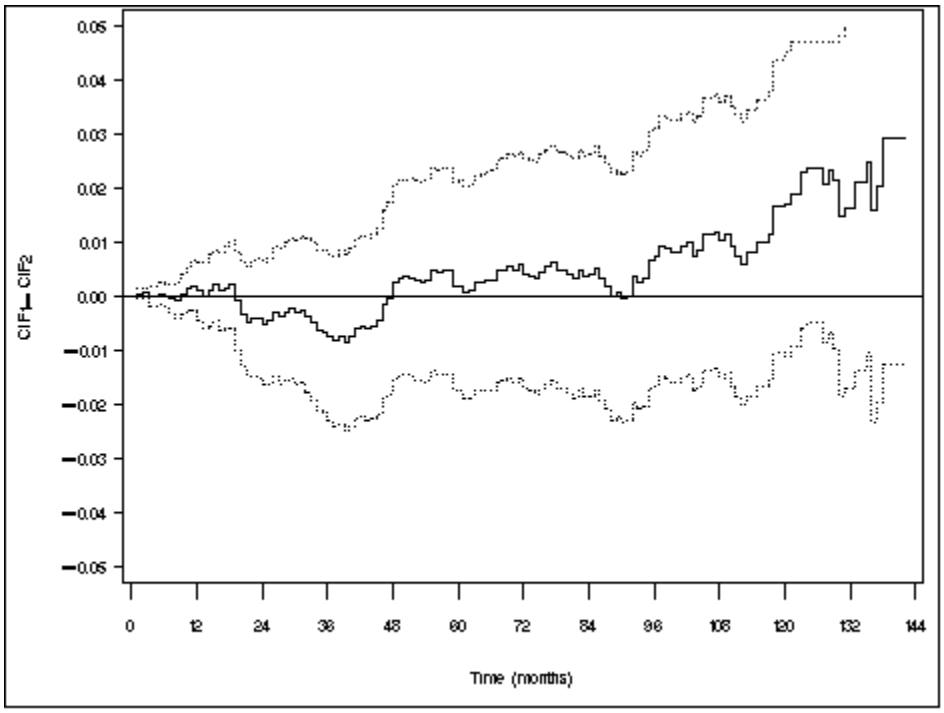


Figure 7 : Direct adjusted cumulative incidence functions of mastectomy and BCS on vascular disease

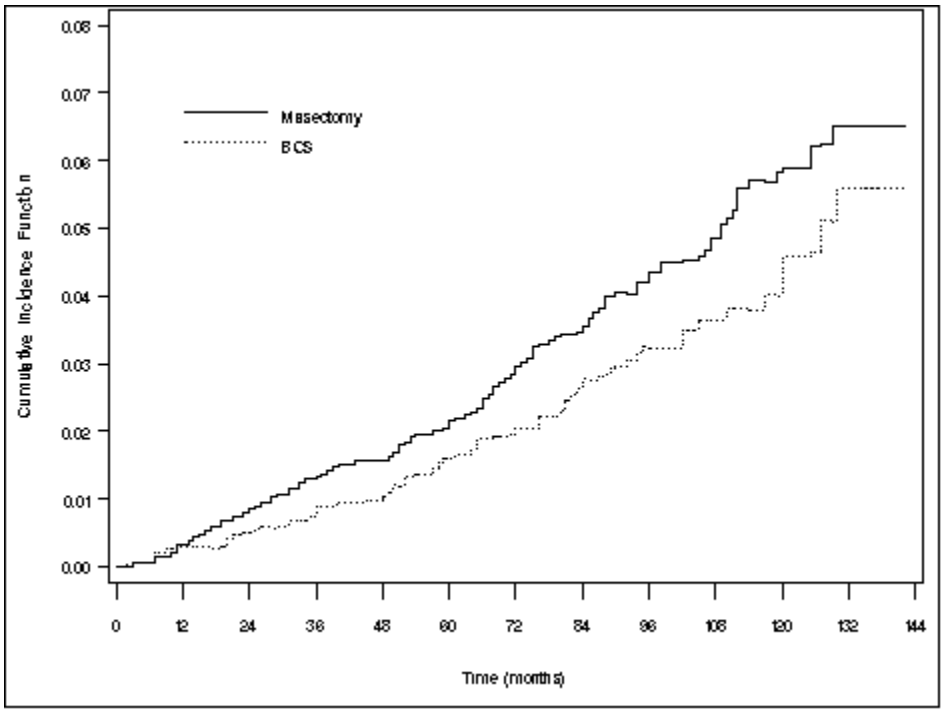


Figure 8: Difference of direct adjusted cumulative incidence functions of mastectomy and BCS on vascular disease

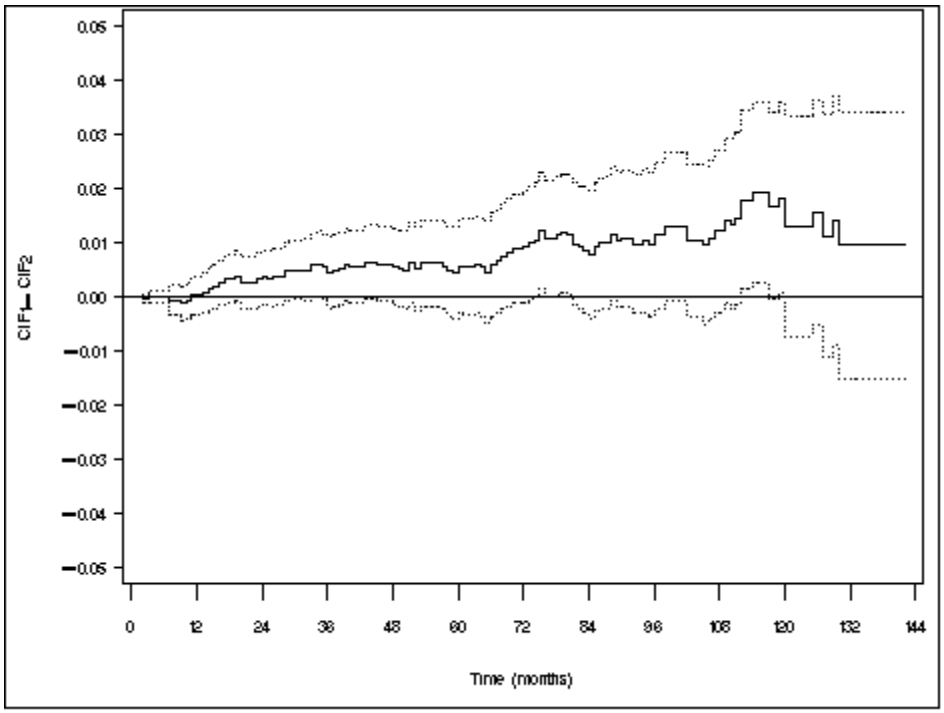


Figure 9 : Direct adjusted cumulative incidence function of mastectomy and BCS on other cause

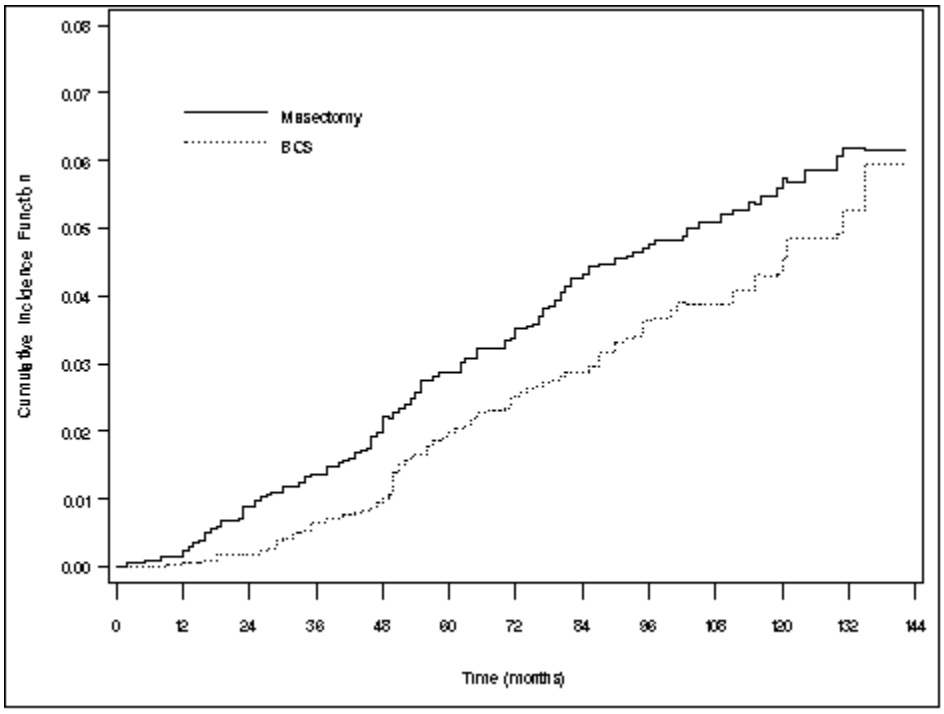


Figure 10 : Difference of direct adjusted cumulative incidence functions of mastectomy and BCS on other cause

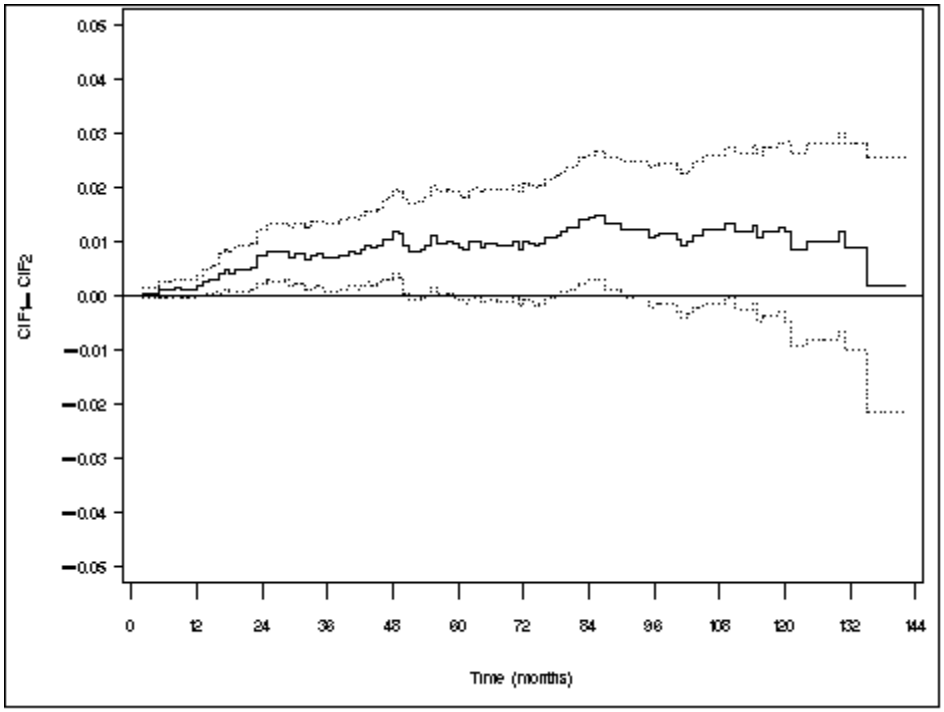


Figure 11 : Direct adjusted cumulative hazards of mastectomy and BCS on cancer

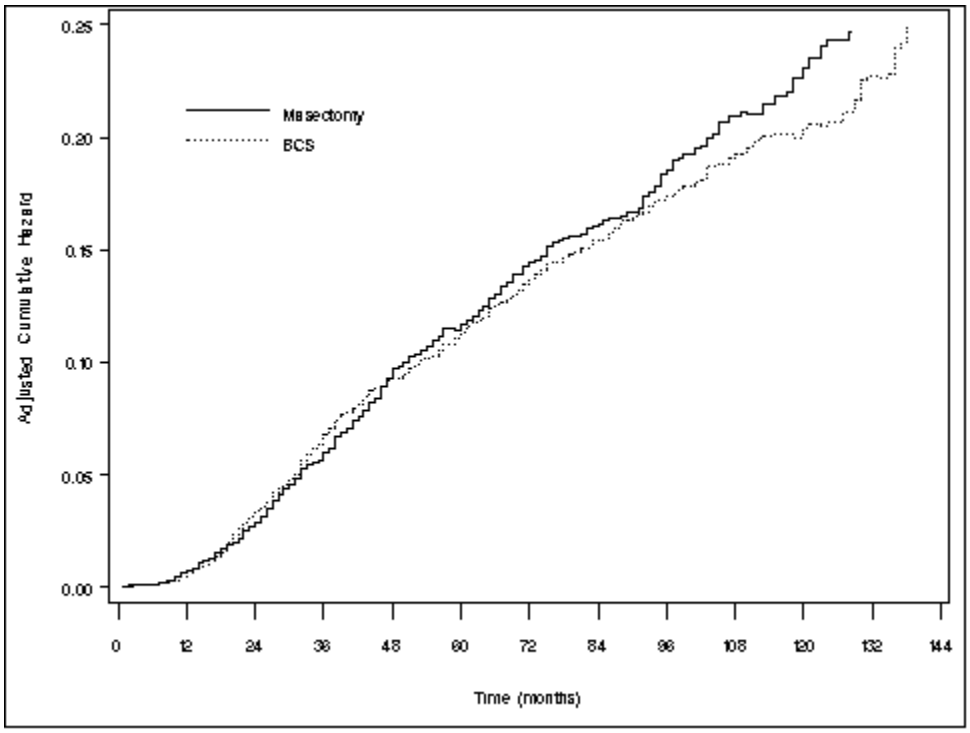


Figure 12: Direct adjusted cumulative hazards of mastectomy and BCS on vascular disease

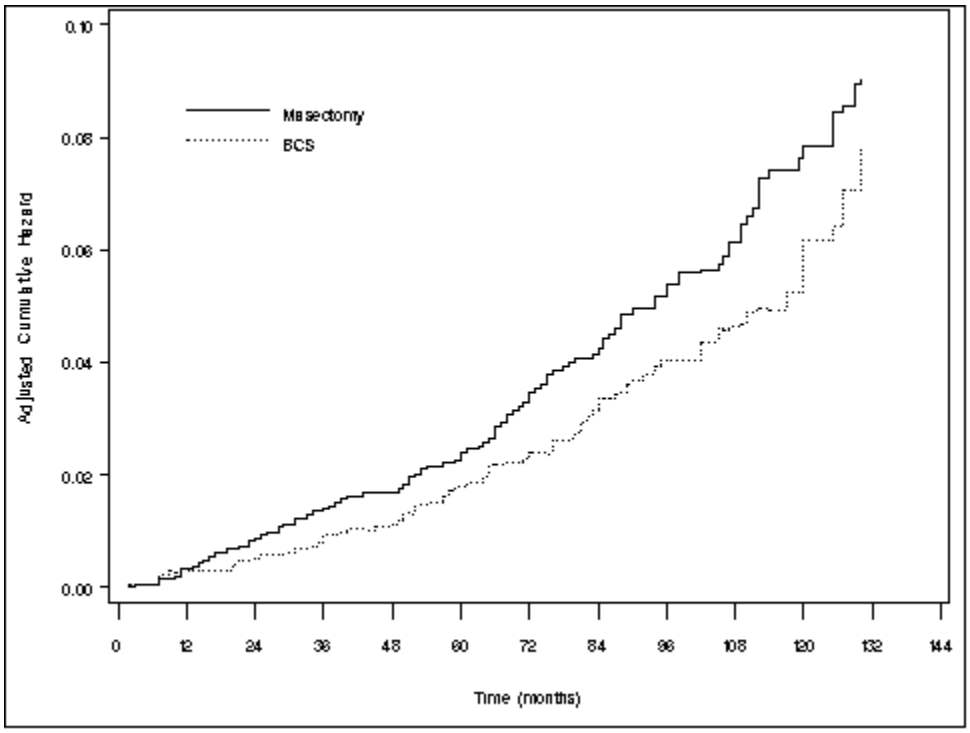


Figure 13 : Direct adjusted cumulative hazards of mastectomy and BCS on other cause

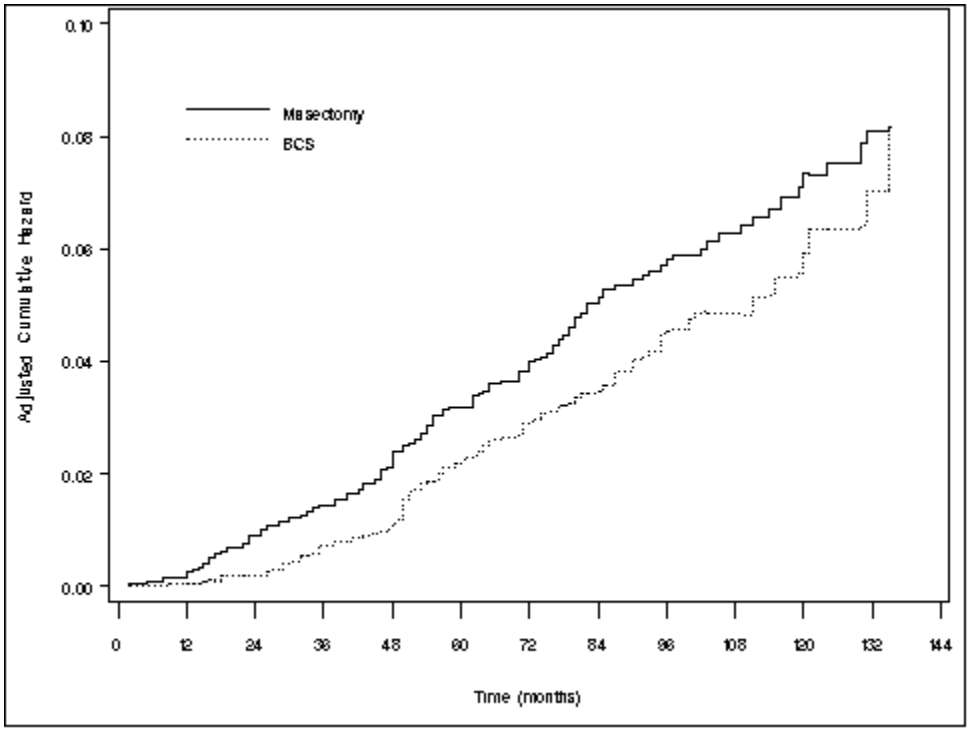
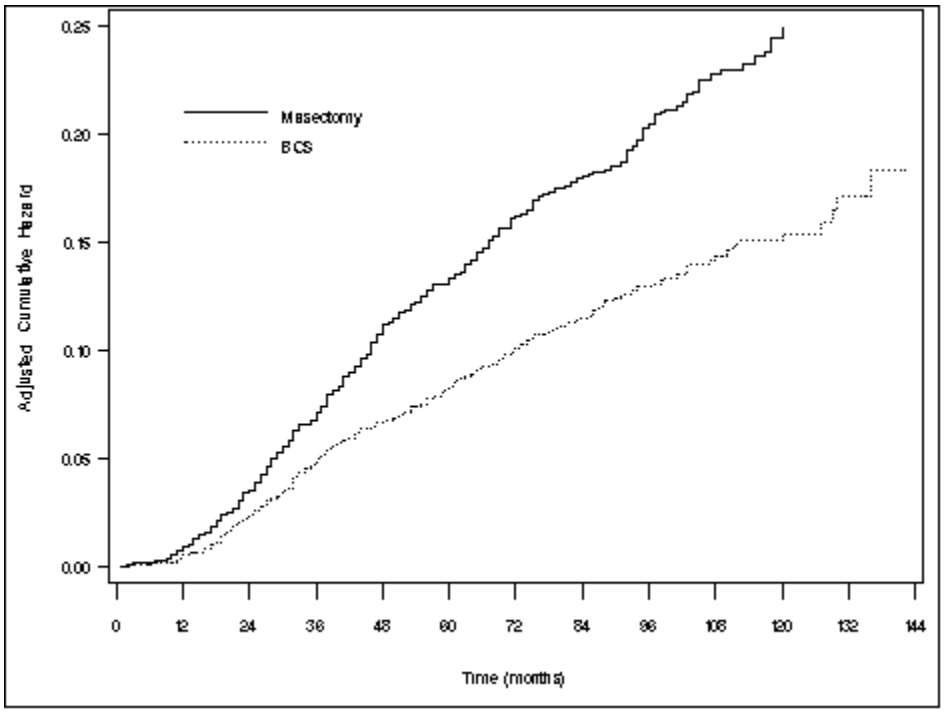


Figure 14: Direct adjusted all-cause cumulative hazard s of mastectomy and BCS



CHAPTER V. CONCLUSION

In this thesis, our goal is to explore the application of Aalen's additive hazards model in analyzing failure time data. Analytical methods implemented in the thesis were motivated by the advantages of Aalen's model: 1. Aalen's model can naturally model influence of a risk factor that changes over time, and does not require constant multiplicative or additive effect assumption for the results to be meaningful. 2. With competing risks data, Aalen's model can be utilized to model all cause-specific hazards, and an additive model still holds for the all-cause hazard.

Based on Aalen's model, we have adopted the direct adjustment method to make a survival curve of a treatment. The direct adjustment method fits Aalen's model very well. The relevant direct adjusted survival curves of different treatments do not need to follow certain patterns required by the parametric or semi-parametric models. This is clearly illustrated by the stem cell transplant example given in Section 4.1. We have plotted the direct adjusted survival curves of three types of transplant in Figure 1. The curve of allogeneic transplant is very different from the curves of the other two types of transplant within the first three years. However, all three curves become close after five years. Treatment comparison at a given time point have been implemented by making a confidence interval for the difference between two direct adjusted survival probabilities. In the stem cell transplant example, we have illustrated that the magnitude of treatment effect changes at different time points.

With competing risks data, we suggest to fit Aalen's model for each cause-specific hazard, and the all-cause hazard, which is the sum of all cause-specific hazard, can be time-varying. Direct adjustment has been also adopted to make the cumulative incidence curve of a treatment, due to one type of failure. Treatment effect at a time, on one type of failure, has been assessed by comparing two direct adjusted cumulative incidence probabilities. In the breast

cancer example, we have compared the survival outcome of two types of surgery, mastectomy versus BCS, with respect to three types of failures, cancer, vascular disease, and other cause. Figures 5 to 10 show that mastectomy yields higher cumulative failures than BCS due to vascular disease or other cause for the majority of the time, while these two types of surgery are quite similar in terms of cancer failures. For comparison, we have made the direct adjusted cumulative hazards of different treatments. The cumulative hazard of a treatment on one type of failure is meaningful only if one hypothetically removes other competing risks.

In this thesis, we have illustrated flexibility of Aalen's model in analyzing failure time data, especially in the context of competing risks. Compared to the commonly used semi-parametric models like Cox (1972) or Lin and Ying's additive model (1994), the power is reduced with Aalen's model to compensate for the complete time-varying effect of all covariates. Various graphical diagnoses have been proposed to show the goodness of fit of a semi-parametric model. It is interesting to further investigate whether a certain criterion can be developed to direct one to adopt a much simple model like Cox or Lin and Ying's model.

Estimation of Aalen's model has been improved by Huffer and McKeague (1991). They proposed a weight least square estimator, which has been show to be more efficient. An extension of current work is to adopt the estimation approach given by Huffer and McKeague. Another important application is to develop robust inference approach to adjust for random effect, also known as frailty, which often appear in failure time data.

REFERENCES

Aalen, O.O. (1989). A linear regression model for the analysis of life times. *Statistics in Medicine* 8, 907-925.

Aalen, O.O. (1993). Further results on the non-parametric linear regression model in survival analysis. *Statistics in Medicine* 12, 1569-1588.

Amato, D.A. (1988). A generalized Kaplan-Meier estimator for heterogeneous populations. *Communications in Statistics-Theory and Methods* 17, 263-286.

Arjas, E. (1988). A graphical method for assessing goodness of fit in cox's proportional hazards model. Journal of the American Statistical Association 83, 204-212.

Besien et al.(2003). A comparison of allogenic and autologous bone marrow transplantation for lymphoblastic lymphoma. *Blood* 101, 2478-2482.

Bretagnolle, J. and Huber-Carol, C. (1988). Effects of omitting covariates in Cox's model for survival data. *Scandinavian Journal of Statistics* 15, 125-138.

Chang, I., Gelman, R. and Pagano, M. (1982). Corrected group prognostic curves and summary statistics. *Journal of Chronic Diseases* 35, 669-674.

Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* 34, 187-220.

Gail, M. H. and Byar, D. P. (1986). Variance calculations for direct adjusted survival curves with applications to testing for no treatment effect. *Biometric Journal* 28, 587-599.

Huffer,W.F., McKeague, I.W. (1991). Weighted least squares estimation for Aalen's additive risk model. *Journal of American Statistical Association* 86, 114-129.

Klein, J.P. and Moescberger, M.L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. 2nd Edition. Springer-Verlag Inc.

Klein, J.P. (2005). Modelling competing risks in cancer studies. *Statistics in Medicine* 25, 1015-1034.

Lee, J.Y., Soong, S.J., Shaw, H.M., Balch, C.M., McCarty, W.H. and Urist, M.M.(1992). Predicting survival and recurrence in localized melanoma: a multivariate approach. *World Journal of Surgery* 16, 191-195.

Lin, D.Y., Ying, Z. (1994). Semiparametric Analysis of the Additive Risk Model. *Biometrika* 81, 61-71.

Makuch, R.W. (1982). Adjusted survival curve estimation using covariates. *Journal of Chron. Disease* 35, 437-444.

Martinussen, T. and Scheike, T.H. (2006). *Dynamic Regression Models for Survival Data.* Springer.

McKeague, I.W. and Sasieni, P.D. (1994). A partly parametric additive risk model. *Biometrika* 81, 501-514.

Neuberger J., Altman, D.G., Christiensen, E., Tygstrup, N. and Williams, R. (1986). Use of prognostic index in evaluation of linear transplantation for primary biliary cirrhosis. *Transplantation* 41, 713-713.

Strutles, C.A. and Kalbfleisch, J.D. (1986). Missepecified proportional hazards models. *Biometrika* 73, 363-369.

Zhang, X., Lobeiza , F.R., Zhang, M.J. and Klein, J.P. (2007). A SAS macro for estimation of direct adjusted survival curves based on a stratified cox regression model. *Computer Methods and Programs in Biomedicine* 88, 95-101.