USE OF PSEUDO-OBSERVATIONS IN THE GOODNESS-OF-FIT TEST FOR GRAY'S TIME-VARYING COEFFICIENTS MODEL

by

Hyung-joo Kang

M.S in Statistics, Michigan State University, 2003

B.S in Mathmatics, Seoul Women's University, Seoul, Korea, 2000

Submitted to the Graduate Faculty of

the Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

Hyung-joo Kang

It was defended on

July 14, 2011

and approved by

Chung-Chou H. Chang, Ph.D, Associate Professor, Departments of Medicine and Biostatistics, School of Medicine and Graduate School of Public Health, University of Pittsburgh

Jong-Hyeon Jeong, Ph.D, Associate Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Lan Kong Ph.D, Assistant Professor, Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh

Susan M. Sereika, Ph.D, Associate Professor, Departments of Health and Community Systems, Biostatistics, and Epidemiology, School of Nursing and Graduate School of Public Health, University of Pittsburgh

Thesis Advisor: Chung-Chou H. Chang, Ph.D, Associate Professor, Departments of Medicine and Biostatistics, School of Medicine and Graduate School of Public Health, University of Pittsburgh

USE OF PSEUDO-OBSERVATIONS IN THE GOODNESS-OF-FIT TEST FOR GRAY'S TIME-VARYING COEFFICIENTS MODEL

Hyung-joo Kang, M.S.

University of Pittsburgh, 2011

Survival analysis has been used to estimate underlying survival or failure probabilities and to estimate the effects of covariates on survival times. The Cox proportional hazards regression model is the most commonly used approach. However, in practical situations, the assumption of proportional hazards (PH) is often violated. The assumption does not hold, for example, in the presence of the time-varying effect of a covariate. Several methods have been proposed to estimate this time-varying effect via a time-varying coefficient. The Gray time-varying coefficients model (TVC) is an extension of the Cox PH model that employs penalized spline functions to estimate time-varying coefficients. Currently, there is no method available to assess the overall goodness-of-fit for the Gray TVC model. In this study, we propose a method based on pseudo-observations. By using pseudo-observations, we are able to calculate residuals for all individuals at all time points. This avoids concerns with the presence of censoring and allows us to apply the residual plots used in general linear regression models to assess the overall goodness of fit for censored survival regression models. Perme and Andersen used the pseudo-observations method to assess the fit for the Cox PH model. We extend their method to assess the fit for the Gray TVC model and illustrate how we applied this approach to assess the fit for a model that predicts posttransplant survival probability among children who were under the age of 12 years, had end-stage liver disease, and underwent liver transplantation between January 2005 and June 2010.

The method has significant public health impact. The Cox PH model is the most cited regression method in medical research. When data violate the PH assumption, The Gray TVC model or an alternative should be used in order to obtain unbiased estimates on survival function and give correct inference on the relationship between potential covariates and survival. The proposed goodness-of-fit test offers a tool to investigate how well the model fits the data. If results show a lack of fit, further modification for the model is necessary in order to obtain more accurate estimates.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to offer my sincere gratitude to my advisor, Dr. Chung-Chou H. Chang, for her encouragement, guidance, patience, time and invaluable input throughout the preparation of this work with patient. I would also like to thank the committee members, Dr. Jong-Hyeon Jeong, Dr. Lan Kong, and Dr.Susan M. Sereika for their valuable comments and suggestions. I appreciate the feedbacks from all of them. I also would like to thank my supervisors, Dr. Judith Erlen and Dr. Leslie Hoffman for their finalcial support and chance for great experience as a biostatistician. Finally, I would like to thank my family and friends for their love, encouragement and support.

1.0 INTRODUCTION

Survival analysis has been used to estimate underlying survival or failure probabilities and to estimate the effects of covariates on survival times. The most commonly used approach is the Cox proportional hazards (PH) regression model [18]. In Cox PH model, the regression coefficients are assumed constant over time. However, in practical situations, the assumption of PH is often violated. The assumption does not hold, for example, in the presence of the time-varying effect of a covariate. If covariate effects vary over time, alternative models which do not necessarily assume proportionality are needed.

Various graphical and numerical methods of goodness-of-fit for checking the Cox model have been introduced. Kalbfleisch and Prentice [41] proposed a graphical method to assess model checking by simply plotting of the log-(-log) survival function vs. time (or log time) for differnt levels of a covariate. Also, a number of graphical methods based on different residuals have been suggested. Kay [44], Andersen [1], and Crowley and Storer [20] proposed methods based on generalized residuals [19]. Schoenfeld [81] and Lin and Wei [54] introduced the methods plotting Schoenfeld residual against time to event. Wei [91] and Therneau, Grambsch, and Fleming [86] suggested methods based on martingale residual. Arjas [9] suggested a method based on comparing between observed and expected failure frequencies. Graphical methods can help decide whether the model fits well; however, the interpretation may be arbitrary without formal significant tests. Numerical methods to examine the adequacy of the fitted model based on Cox regression model also have been studied. In 1972, Cox [18] suggested a model checking method using time-varying covariates. Schoenfeld [80], Moreau, O'Quigley, and Mesbah [64], and Moreau, O'Quigley, and Lellouch [63] constructed goodness-of-fit tests by partitioning subjects based on mutually exclusive regions from the product space of the time and covariates.

Similar approaches were proposed by Arjas [9], Andersen et al. [2], and Grambsch and Therneau [30]. Wei [91] constructed the tests using two-sample hazards ratio. Gill and Schumacher [27] proposed tests of the methods via generalized rank estimators of the relative risk. Lin [53] proposed an analytic test to check the proportional hazards assumption. Horowitz and Neumann [39] and Lin and Wei [54] constructed global goodness-of-fit test. Other numerical approaches of goodness-of-fit for Cox model have been studied by Andersen and Gill [4], Moreau et al [64], Ciampi and Etezadi-Amoli [17], Nagelkerke, Oosting and Hart [67], and O'Quigley and Pessione [69]. Gray [34] and Hess [38] defined a family of alternative functions based on unspecified smoothing functions to the covariate or coefficients. Their approach allows testing the global validity of the model as any specified time-dependent alternative. A simulation study conducted by Ezzeddine [23] showed that Lin's [53] and Gray's [34] tests are the most powerful for assessing the proportionality assumption.

Several methods have been proposed to estimate time-varying effect via a time-varying coefficient. The time-varying coefficient models were considered previously by Gamerman and West [26], Zucker and Karr [93], and Gray [33] to model with time-varying effects and to test PH assumption. In these models products of a covariate and spline functions of time are used to fit the model. Zucker and Karr [93], Andersen and Gill [4], Gill [28], and O'Sullivan [70] introduced nonparametric estimation in the Cox model which is based on penalized likelihood. Varying-coefficient (VC) models were also introduced by Hastie and Tibshirani [37] for time-varying coefficient models. Also, the extended Cox model with smoothing spline was proposed by Sleeper and Harrington [83] and Gray [33]. The Sleeper and Harrington approach uses the partial likelihood, while Gray approach uses the penalized partial likelihood to estimate the parameters. Li, Klein, and Moeschberger [51], and Persson [74] proposed the stratified Cox regression model by stratification of the covariate which violates PH assumption. Pettitt and Daud [75] and Fisher and Lin [24] proposed extending Cox regression model by including time-dependent covariate and product of time and time-dependent covariate.

In 1992 Gray [33] introduced a time-varying coefficients model. This model is an extension of Cox's PH model that employs penalized spline functions to estimate time-varying coefficients. The basic concept of Gray's time-varying coefficients model is to use smoothing spline functions at modified number of time knots, and then to estimate the parameters of the model by maximizing the penalized partial likelihood. Advantage of Gray's piecewise-constant time-varying model is that it is computationally less complex.

Currently, there is no method available to assess the overall goodness-of-fit for Gray's model. In this study, we propose a goodness-of-fit method based on pseudo-observations [7]. Pseudo-observations can be calculated for all individual at all time points. We will then use the pseudo-observation to replace incomplete observed outcome, censored observation, to create a pseudo sample. Graphical analysis, especially pseudo residual plots, can be conducted using the same logic of general linear models [73] to assess the model fit.

In Section 2, terminologies and notations used in this paper are defined. Also, Cox proportional hazards model, Gray's piecewise-constant time-varying coefficient model and pseudo-observation are defined in this section. Finally, in Section 2, goodness-of-fit tests using pseudo residuals based on survival model are described. In Section 3, simulations are presented to see how well Gray's piecewise-constant time-varying coefficient model fits the data set. In Section 4, we illustrate how we applied this approach to assess the fit for a model that predicts posttransplant survival probability among children who were under the age of 12 years, had end-stage liver disease, and underwent liver transplantation between January 2005 and June 2010. The conclusion and limitations are discussed in Section 5.

2.0 DESCRIPTION OF STATISTICAL METHODS

2.1 TERMINOLOGY & NOTATIONS

The right censored survival data of sample size n contains (T_i, δ_i, Z_{ij}), where i ($i = 1, \ldots, n$) denotes the i^{th} subject, and j ($j = 1, \ldots, p$) denotes the j^{th} risk factor:

X_i: the survival time, the failure time or the time to event for the i^{th} subject, $X_i > 0$.

C_i: the censoring time for the i^{th} subject, $C_i > 0$.

$T_i = min(X_i, C_i)$, which is the observed time on study for the i^{th} subject.

δ_i: the event indicator for the i^{th} subject

$$\delta_i = \begin{cases} 1, & \text{if an individual } i\text{has experienced the event}(X_i \leq C_i) \\ 0, & \text{if an individual } i\text{is censored}(X_i > C_i) \end{cases}$$

$Z_i(t) = (Z_{i1}(t), ..., Z_{ip}(t))$: the vector of covariates or vector of risk factors for the i^{th} subject at time t, which may time-constant or time-varying covariates. In this study, we focus on time-constant covariate. Therefore, we denote Z as $Z_i = (Z_{i1}, ..., Z_{ip})$.

$\beta(t) = (\beta_1(t), ..., \beta_p(t))$: the vector of regression coefficient at time t, which may time-constant (Cox PH model) or time-varying (Gray's time-varying model).

A right-censoring at time t implies that the event of interest has not yet occurred before or at time t. The failure time X_i and the censoring time C_i for the i^{th} subject are assumed to be independent. Also, censoring times is assumed to be independent with covariates in the fitted model.

2.2 COX PROPORTIONAL HAZARDS MODEL

Cox [18] proposed a proportional hazards model to estimate underlying survival or failure probabilities and to estimate the effects of covariates on survival times via the hazard function $h(t|Z)$. The hazard function is formed as

$$h(t|Z) = h_0(t)\exp(\beta' Z)\,, \tag{2.1}$$

where $h_0(t)$ is a baseline hazard which is unknown and unspecified non-negative function, t is time to occurrence of some event, $Z_i = (Z_{i1}, ..., Z_{ip})$ is the vector of covariates or vector of risk factors, and $\beta' = (\beta_1, \ldots, \beta_p)$ is the vector of regression parameters which implies the effects of risk factors. The Cox regression model (2.1) holds two assumptions: the proportional hazards (PH) assumption and the linearity assumption for the continuous covariates. The Cox PH assumption means that the regression coefficient β is assumed to be constant over time.

Comparing hazard rate of an individual with a set of risk factors Z to hazard rate of any other individual with a set of risk factors Z^*, which is called as the hazard ratio (HR), is defined as

$$HR \;=\; \frac{h_0(t)\exp(\beta' Z)}{h_0(t)\exp(\beta' Z^*)} = \exp(\beta'(Z - Z^*)) \tag{2.2}$$

Since (2.2) is independent of time t, the hazard ratio of an individual with risk factors Z_1 experiencing the event comparing to any other individual with risk factors Z_2 experiencing the event is also independent of time t, i.e. proportional. When Z is a binary variable, where $Z = 1$ if an individual received the treatment, and $Z = 0$ if an individual did not received the treatment, to evaluate the treatment effect, the hazard ratio of $Z = 1$ to $Z^* = 0$ is computed as

$$HR \;=\; \exp(\beta'(Z - Z^*)) = \exp(\beta') \tag{2.3}$$

It means that the hazard ratio of an individual who received treatment and experiences the event compared to an individual who did not receive treatment and experienced the event, $\exp(\beta)$, is constant over time. The log-hazard ratio is

$$\ln(HR) = \ln(\frac{h_0(t)\exp(\beta' Z)}{h_0(t)\exp(\beta' Z^*)}) = \ln(\exp(\beta'(Z - Z^*))) = \beta'(Z - Z^*)\,. \quad (2.4)$$

As mentioned in Section 2.1, the data set of a survival analysis contains $(T_i,\ \delta_i,\ Z_{ij})$, where i $(i = 1,\ \ldots,\ n)$ denotes the i^{th} subject, and j $(j = 1,\ \ldots,\ p)$ denotes the j^{th} risk factor. Also, the failure time X_i and the censoring time C_i for the i^{th} subject are assumed to be independent. Then, the parameter β in Cox PH model is estimated by maximizing the partial likelihood function. There are two situations; partial likelihood for distinct failure time and partial likelihood for tied failure time.

For the first situation, it is assumed that all subjects in the study have D distinct event time, $t_1 < \ldots < t_D$. The partial likelihood function [18] to estimate parameter β in Cox PH model is given by

$$L(\beta) \;=\; \prod_{k=1}^{D} \left\{ \frac{\exp(\beta' Z_{(k)})}{\sum_{i \in R(t_k)} \exp(\beta' Z_i)} \right\}\,, \quad (2.5)$$

where $R(t_k)$ implies the set of those subjects who are still at risk just prior to time t_k, and $Z_{(k)}$ is set of the covariates related with the individual whose failure time is t_k.

The second situation is the case of existing tied failure time. The partial likelihood function proposed by Breslow [15] to estimate parameter β in Cox PH model for existing tied failure time is given by

$$L(\beta) \;=\; \prod_{k=1}^{D} \left\{ \frac{\exp(\beta' S_k)}{\left[\sum_{i \in R(t_k)} \exp(\beta' Z_i)\right]^{\delta_k}} \right\}\,, \quad (2.6)$$

where δ_i is the number of failure times equal to t_k, and S_k is the sum of vector Z_i for all subjects who experienced the event at t_k.

To obtain an estimated survival function, we need to estimate the cumulative baseline hazard function, $H_0(t) = \int_0^t h_0(s)ds$. In this study, Breslow's estimator [14] is used:

$$\hat{H}_0(t) = \sum_{t_k \leq t} \frac{1}{\sum_{j \in R(t_k)} \exp(\beta' Z_i)} \tag{2.7}$$

The cumulative hazard function can be calculated by integrating hazard function (2.2) over the range of time t:

$$H(t|Z) = \int_0^t h_0(s) \exp(\beta' Z^*) ds \tag{2.8}$$

The relation of hazard function and survival function is

$$\begin{aligned} S(t|Z = Z) &= \exp(-H(t|Z = Z)) \\ &= \exp[-\int_0^t h_0(s) \exp(\beta' Z) ds] \\ &= \exp[-\exp(\beta' Z) \int_0^t h_0(s) ds] \\ &= \exp[-H_o(t)]^{\exp(\beta' Z)} \\ &= [S_0(t)]^{\exp(\beta' Z)} \end{aligned} \tag{2.9}$$

where $S_0(t)$ is the baseline survival function. The estimated survival function at time t, given Z=z^* , then can be obtained by (2.9)

$$\begin{aligned} \hat{S}(t|Z = z^*) &= \exp(-\hat{H}(t|Z = z^*)) \\ &= \exp[-\int_0^t \hat{h}_0(s) \exp(\hat{\beta}' z^*) ds] \\ &= \exp[-\exp(\hat{\beta}' z^*) \int_0^t \hat{h}_0(s) ds] \\ &= \exp[-\hat{H}_0(t)] \exp^{\exp(\hat{\beta}' z^*)} \\ &= [\hat{S}_0(t)]^{\exp(\hat{\beta}' z^*)} \end{aligned} \tag{2.10}$$

where $\hat{S}_0(t)$ is the estimated baseline survival function. $\hat{H}_0(t)$ is calculated by Breslow's estimator [14] and $\hat{\beta}$ is calculated by maximizing the partial likelihood.

2.3 GRAY'S PIECEWISE-CONSTANT TIME-VARYING COEFFICIENTS MODEL

In many cases of clinical trial, the effects of risk factors may change over time. In this case, the Cox proportional hazards assumption is violated, and the Cox PH regression model is not valid any more. Therefore, other approaches to deal with a time-varying coefficient, such as extended Cox model or Gray's time-varying coefficients model, are required. The time-varying coefficients models were proposed by Gray [33] to model with time-varying effects over time and to test PH assumption. Gray's time-varying model allows violation of the proportional hazards assumption via piecewise-constant time-varying coefficient using penalized B-splines. The advantage of Gray's model is its flexibility because the proportional hazards assumption is assumed to hold only for the each time interval.

Hazard function using Gray's time-varying coefficient model is formed as

$$h(t|Z) = h_0(t)\exp(\beta(t)'Z)\,, \tag{2.11}$$

where, $h_0(t)$ is a baseline hazard, which is unknown and unspecified non-negative function, t is time to occurrence of some event, $Z_i = (Z_{i1}, ..., Z_{ij})$ is the vector of covariates or vector of risk factors, and $\beta(t)' = (\beta_1(t), \beta_2(t), ..., \beta_p(t))$ is the vector of regression parameters which implies the effects of risk factors.

Comparing the hazard rate of an individual with a set of risk factors Z to the hazard rate of any other individual with a set of risk factors Z^*, the hazard ratio is

$$HR \;=\; \frac{h_0(t)\exp(\beta(t)'Z)}{h_0(t)\exp(\beta(t)'Z^*)} = \exp(\beta(t)'(Z - Z^*))\,. \tag{2.12}$$

Since (2.12) is time dependent, the hazard ratio of an individual with risk factors Z_1 experiencing the event comparing to any other individual with risk factors Z_2 experiencing the event is also time dependent, i.e. nonproportional.

When Z is binary variable, say $Z = 1$ if an individual received the treatment, and $Z = 0$ if an individual did not received the treatment, to evaluate treatment effect, the hazard ratio of $Z = 1$ to $Z^* = 0$ is computed as

$$HR = \exp(\beta(t)'(Z - Z^*)) = \exp(\beta(t)') \; . \tag{2.13}$$

It means that the hazard ratio of an individual who received treatment and experiences the event compared to an individual who did not receive treatment and experienced the event, $\exp(\beta(t))$, varies over time. The log-hazard ratio is

$$\ln(HR) \;\; = \;\; \ln(\frac{h_0(t)\exp(\beta'(t)Z)}{h_0(t)\exp(\beta'(t)Z^*)}) = \beta(t)'(Z - Z^*) \; . \tag{2.14}$$

In the vector of regression parameters $\beta(t)' = (\beta_1(t), \beta_2(t), ..., \beta_p(t))$, the function of time-varying coefficients is formed as

$$\beta_j(t) = \sum_k \theta_{jk}\beta_{jk}(t) \; , \tag{2.15}$$

where j $(j = 1, \; ... \; , \; p)$, denotes the j^{th} risk factors in the model, k denotes indexes of the time intervals, $(k = 1, \; ... \; , \; M + 1)$.

In (2.15), $\beta_{jk}(t)$ is called B-spline basis function [21]. The spline with only certain number of knots and with estimated parameters $\hat{\beta}_j(t)$ are used in Gray's piecewise-constant time-varying coefficients model. As mentioned before, in Gray's piecewise-constant time-varying coefficients model, the regression coefficients β_j(t) remain constant on time intervals between the selected time knots, $t \in [\tau_k, \; \tau_{k+1})$, and the regression coefficients β_j(t) are allowed to change at the selected internal time knots, τ_k, where $\tau_0 = 0$, and $\tau_{M+1} = T$, which is maximum observed time. Therefore, the coefficients β_j(t) are right-continuous step function of time, and jumps at each time knot. The time knots are predetermined to be roughly equally spaced on the event scale.

The parameters $\beta_j(t)$ in Gray's piecewise-constant time-varying coefficients model are estimated by maximizing the penalized partial likelihood function [33]. The penalized partial likelihood is combination of the usual log partial likelihood $L(\beta)$ [18] (2.5) and piecewise-constant function with penalty $\frac{1}{2}\lambda_j \sum_{k=2}^{M+1}(\theta_{jk} - \theta_{j,k-1})^2$ [33], where M is number of internal time knots. The penalized partial likelihood function is given by

$$L_p(\beta) = L(\beta) - \frac{1}{2}\lambda_j \sum_{k=2}^{M+1}(\theta_{jk} - \theta_{j,k-1})^2 \,. \tag{2.16}$$

To drive the estimated cumulative hazard $H_0(t) = \int_0^t h_0(s)ds$ in (2.8), Valenta [89] suggested the estimator of cumulative baseline hazard based on Breslow's estimator [14]

$$\hat{H}_{oj}(t) = \int_{[\tau_k,\tau_{k+1})} I(u \le t)\hat{h}_0(u) = \int_{[\tau_k,\tau_{k+1})} I(u \le t)\frac{\sum_{i=1}^n dN_i(u)}{\sum_i Y_i(u)\exp\left\{z_i'\hat{\beta}(u)\right\}} \,, \tag{2.17}$$

which is a part of the cumulative baseline hazard function, $H_0(t)$, on the interval $[\tau_k, \tau_{k+1})$ ($k = 1, \ldots, M$), where $\tau_0 = 0$, $\tau_{M+1} = T$, which is maximum observed time (failure or censored time). $Y_i(t)$ is an indicator for the subject who is still at risk just prior to time t, and $dN_i(t)$ is the change in the process $N(t)$ over a short time interval $[t,\ t + dt]$.

The survival function based on the Gray's piecewise-constant time varying coefficient model is formed as

$$S(t|Z) = \exp[-\int_o^t h_{0j}(s)\exp(\beta_j(s)'Z)ds] \,. \tag{2.18}$$

Since we use fixed internal time knots, the survival function suggested by Valenta [89] is used

$$\hat{S}(t|Z) \;=\; \exp[-\sum_{k=0}^{M}\hat{H}_{0k}(t)\exp(\hat{\beta}_k'Z)] \,, \tag{2.19}$$

where $\hat{\beta}_k = \hat{\beta}(\tau_k)$, which is the estimated regression coefficient at τ_k time knot, k ($k = 0, \ldots, M$) denotes the k^{th} time knot, and M is number of internal time knots. $\hat{H}_0(t)$ is derived by Breslow's estimator [14], and $\hat{\beta}_j$ is calculated by maximizing the penalized partial likelihood [33].

2.4 PSEUDO-OBSERVATIONS

Let $\theta = E(f(X))$, and $\hat{\theta}$ be an unbiased estimator of θ, where $f(\cdot)$ could be any defined function. Given this, Anderesen et al. [7] defined a pseudo-observation for the i^{th} individual as

$$\hat{\theta}_i = n \cdot \hat{\theta} - (n-1) \cdot \hat{\theta}_{-i} \, , \tag{2.20}$$

where $\hat{\theta}_{-i}$ is the "leave-one-out" estimator for θ, that is, the estimate is computed from the sample in which the i^{th} individual is removed. This idea is from the Jackknife methodology [61]. The average of all pseudo-observations is the jackknife estimate of θ [13]. From definition (2.20), pseudo-observations can be calculated for all individuals. The idea of pseudo-observation is to replace incompletely observed $f(X_i)$ by pseudo-observation $\hat{\theta}_i$ to achieve complete data.

For survival data, if there is censoring, the data of event history is incomplete, i.e., the survival times would not be observed for all subjects. Let $X = (X_1, X_2, \ldots, X_n)$ be the vector of failure times for n individuals, where X_i's $(i = 1, \ldots, n)$ are independent and identically distributed. Survival indicator for the i^{th} subject at time t, where t implies time point $(t = t_1, \ldots, t_k)$, is defined by

$$f(x) = f_t(x) = I(X_i > t) \qquad . \tag{2.21}$$

While in principal the survival indicator for the i^{th} subject is observed for all individuals at all time points, in practical it may not be observed for some subjects at some time points because of censoring.

Survival probability at time t is estimated by expectation of survival indicator

$$S(t) = E(I(X_i > t)) = E(f_t(x)) \, ,$$

and then parameter θ is $S(t)$ under (2.20).

Pseudo-observation for the survival function [73] is defined as

$$S_i(t) = n\hat{S}(t) - (n-1)\hat{S}_{-i}(t)\ , \tag{2.22}$$

where $\hat{S}(t)$ is Kaplan-Meier estimator based on the whole sample [42]:

$$\hat{S}(t) = \prod_{u \le t}\left(1 - \frac{dN(u)}{Y(u)}\right)\ , \tag{2.23}$$

and $\hat{S}_{-i}(t)$ is Kaplan-Meier estimator based on the sample omitting the i^{th} observation. Under the assumption of independent censoring, the average of the pseudo-observations for the survival function computed at each time point is close to the estimated survival function based on the Kaplan-Meier estimator $\hat{S}(t)$ and the true value of $S(t)$ [84].

$$\hat{S}(t) = \prod_{u \le t}\left(1 - \frac{dN(u)}{Y(u)}\right) = \frac{1}{n}\sum_{i=1}^{n} S_i(t) \approx S(t)\ . \tag{2.24}$$

If there is no censoring, $S_i(t)$ can be expressed with index function, specified as survival indicator: $S_i(t) = I(X_i > t)$, $S_i(t) = 1$ if individual is still alive at time t, $S_i(t) = 0$ if individual experience the event before time t. Even when there is censoring, the pseudo-observations are defined for all individual and at all time points. Pseudo-observations for each individual are plotted in Figure 1 and Figure 2. In Figure 1 and Figure 2, pseudo-observation plots show the jump at the time of event for the individual. Some values are greater than 1 because omitting individuals at risk at certain time points reduces the risk set.

2.5 GOODNESS-OF-FIT TEST USING PSEUDO-OBSERVATIONS FOR SURVIVAL MODELS

In Cox PH model, (2.1), the regression coefficients are assumed to be constant over time. However, in practical situations, the assumption of proportional hazards (PH) is often violated. Since the Cox regression model estimates parameters based on PH assumption, the Cox PH regression model will be invalid when non-proportionality exists and the estimate will be under or overestimated under the violation of this assumption. To check model fit, goodness-of-fit test is required. Currently, there is no method available to assess the overall goodness-of-fit for Gray's time-varying coefficients model. In this study, we propose a goodness-of-fit method for Gray's model based on pseudo-observations [7]. One way for assessing goodness-of-fit test is to analyze residuals. However, in survival analysis, the plotting may be faced with troubles because of presence of censoring: the survival indicator, $f_t(x) = I(X_i > t)$, is not always observed. As mentioned before, the pseudo-observations for survival function, $S_i(t)$, are observed for all individual and at all time points. The idea is to replace the survival indicator by its pseudo-observation; therefore, using pseudo-observation, standard residual analysis can be conducted regardless of censoring under the same logic of general linear models [73] to assess the model fit.

The idea of pseudo residual is to compare pseudo-observations for survival function, denoted as $S_i(t)$, with predicted values of survival function based on the fitted model, denoted as $\hat{S}(t|Z_i)$. Perme and Andersen [73] defined the raw pseudo residual as

$$\hat{r}_i(t) = S_i(t) - \hat{S}(t|Z_i) \, , \tag{2.25}$$

and defined the standardized pseudo residual as

$$\hat{e}_i(t) = \frac{S_i(t) - \hat{S}(t|Z_i)}{\sqrt{\hat{S}(t|Z_i)[1 - \hat{S}(t|Z_i)}} \, . \tag{2.26}$$

We will use pseudo residual as a graphical diagnostic tool to evaluate model fit for Gray's time-varying coefficients model. Raw pseudo residual for Gray's time-varying model is defined as

$$\hat{r}_{i,Gray}(t) = S_i(t) - \hat{S}_{Gray}(t|Z_i) \, , \tag{2.27}$$

and standardized pseudo residuals for Gray's time-varying model is defined as

$$\hat{e}_{i,Gray}(t) = \frac{S_i(t) - \hat{S}_{Gray}(t|Z_i)}{\sqrt{\hat{S}_{Gray}(t|Z_i)[1 - \hat{S}_{Gray}(t|Z_i)]}} \, . \tag{2.28}$$

Finally, pseudo residuals are plotted against estimated survival function based on fitted model at each time point with smoothed averages. Since many points are overlapped on the residual plots, it is impossible to evaluate the trends by using the residuals alone. Therefore, smoothed average plot along with pseudo residual is necessary. If the model shows a good fit in estimating survival function, the smoothed averages stay around zero. If the model does not fit well, some type of departures and tendencies are shown on the residual plot.

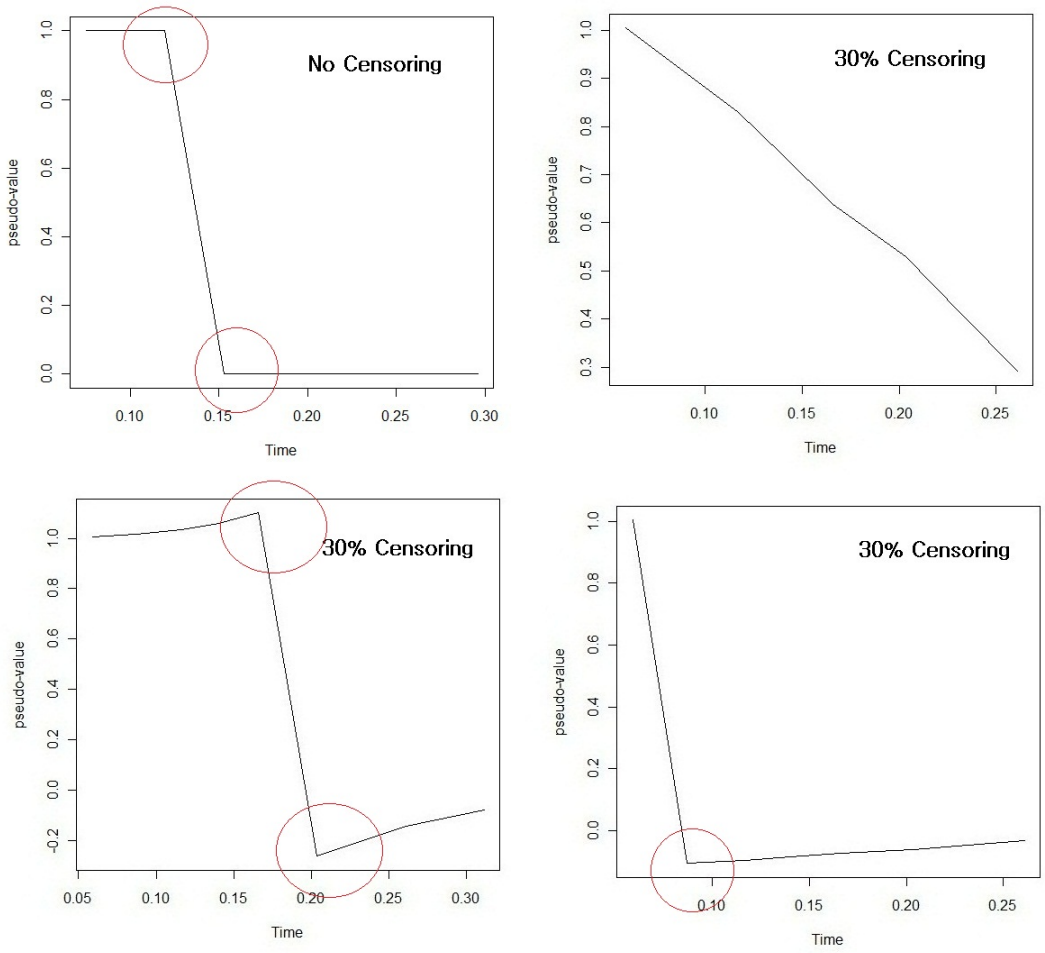

Figure 1: The pseudo-observations for an individual

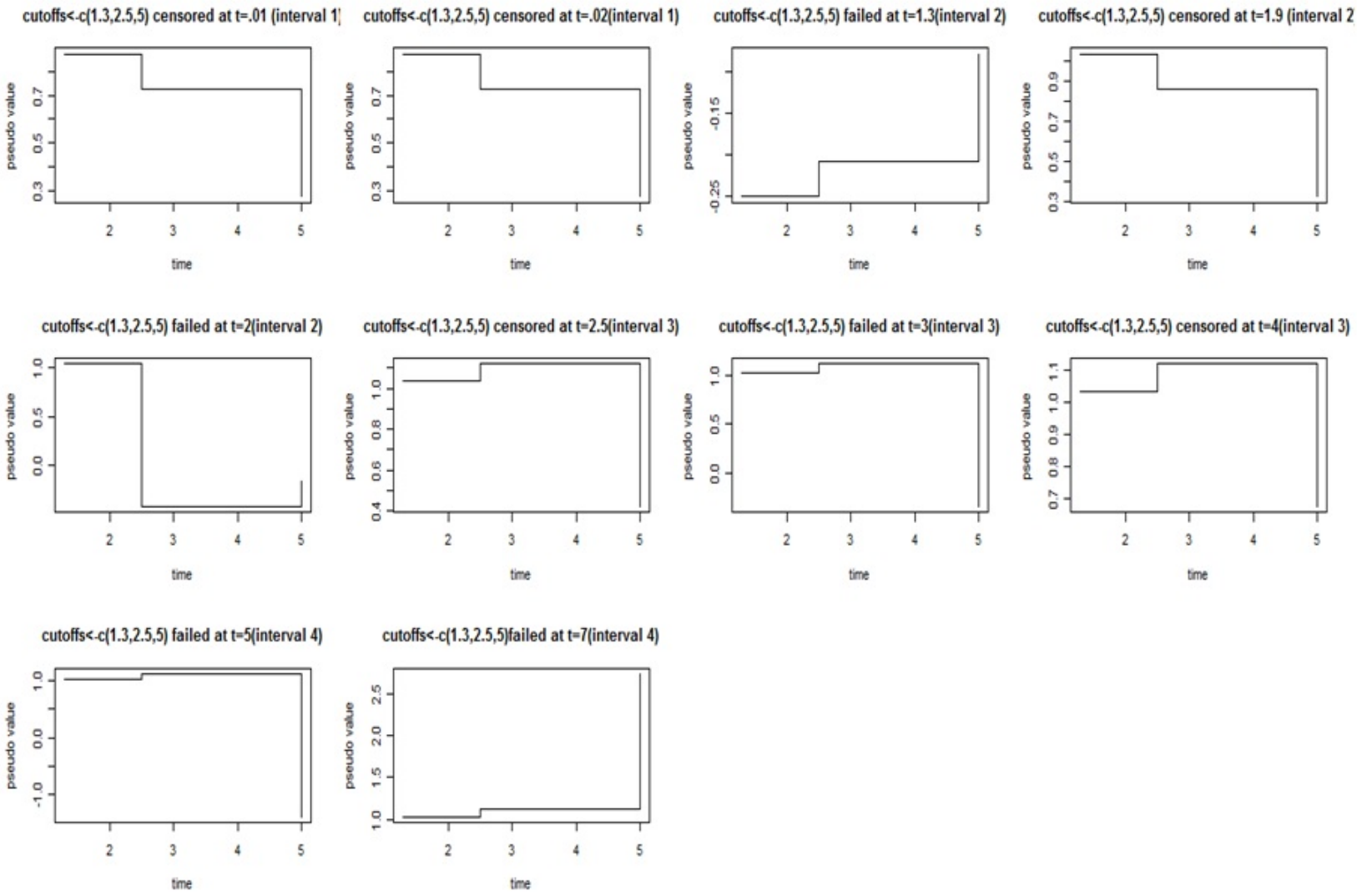

Figure 2: The pseudo-observations when there exists any censoring

3.0 SIMULATION STUDY

3.1 OVERVIEW METHODS

Simulation studies were performed under scenarios of violating the proportional hazards assumption, and the experiment was repeated 1000 times with samples of size 500. Both complete and right-censored survival times were generated in the following settings.

Let $\beta(t)$ be the time-varying effect of a covariate Z, where $f(Z)$ is assumed to be linear. Hazard function for the time-varying model (2.11) is defined as

$$h(t|Z) = h_0(t)\exp(\beta(t)Z)\,,$$

where $h_0(t) = \lambda$, which is a constant baseline hazard function, and $\beta(t) = \theta t$, which is a time-varying regression coefficient.

The corresponding cumulative hazard function becomes

$$H(t|Z) = \int_0^t \lambda\exp(\theta x Z)dx = \frac{\lambda}{\theta Z}(\exp(\theta t Z) - 1)\,, \tag{3.1}$$

and the survival function has the form of

$$S(t|Z) = \exp(-H(t|Z)) = \exp(-\frac{\lambda}{\theta Z}(\exp(\theta t Z) - 1)\,, \tag{3.2}$$

and the distribution function of Gray's model is

$$F(t|Z) = U = 1 - S(t|Z) = 1 - \exp(-\frac{\lambda}{\theta Z}(\exp(\theta t Z) - 1))\,, \tag{3.3}$$

where $U \sim$ U[0,1].

Let X be the survival time,

$$U = 1 - \exp(-\frac{\lambda}{\theta Z}(\exp(\theta X Z) - 1)) \ ,$$

where $U \sim$ U[0,1].

To generate survival time X,

$$F^{-1}(t|Z) = X = [\ln(-\frac{(\theta Z \ln(1-U)}{\lambda} + 1)/(\theta z)] \ , \tag{3.4}$$

where U is a random variable following uniform distribution on the interval from 0 to 1, U[0,1].

In previous section, hazard ratio is formed as 2.12,

$$HR \ = \ \frac{h_0(t)\exp(\beta(t)'Z)}{h_0(t)\exp(\beta(t)'Z^*)} = \exp(\beta(t)'(Z - Z^*))$$

and the log hazard ratio is expressed as 2.14,

$$\ln(HR) \ = \ \ln(\frac{h_0(t)\exp(\beta'(t)Z)}{h_0(t)\exp(\beta'(t)Z^*)}) = \beta(t)'(Z - Z^*) \ .$$

The log hazard ratio is a linear form of $\beta(t)$. Therefore, the variety of value of $\beta(t)$ over time can be checked using log hazard ratio plot vs. time.

Simulations were conducted with varying percentages of censoring with fixed $\lambda = 1$, and $\theta = 2$. Under this simulation setting, true hazard function is formed as

$$h(t|Z) \ = \ h_0(t)\exp(2tZ)$$

In this study, we generated (T_i, δ_i, Z_i), where i ($i = 1, \ldots, n$) denoted the i^{th} subject. X_i is the survival time for the i^{th} subject, which was generated by (3.5). C_i is the censoring time for the i^{th} subject, which was generated following exponential distribution with parameter r. Censoring rate is changed with different choice of r. T_i is the observed time on study for the i^{th} subject, which is obtained by $T_i = min(X_i, C_i)$, .

δ_i is the event indicator for the i^{th} subject :

$$\delta_i = \begin{cases} 1, & \text{if the individual i has experienced the event}(X_i \leq C_i) \\ 0, & \text{if the individual i is censored}(X_i > C_i) \end{cases}$$

For simplicity, only one risk factor, Z, was used in this simulation.

3.2 SIMULATION I: COMPARING AVERAGE OF PSEUDO-OBSERVATIONS TO ESTIMATED SURVIVAL FUNCTION BASED ON KAPLAN-MEIER ESTIMATOR

3.2.1 Generate Data Set and Overview Methods

The main goal of simulation I is to show that average of the pseudo-observations is close to the Kaplan-Meier estimator. The risk factor Z_i for the i^{th} subject is a binary variable ($Z_i = 0$ or $Z_i = 1$), which is the time-varying coefficient covariate. Let $h_0(t) = \lambda = 1$, which is a constant baseline hazard function, and $\theta = 2$ and $b(t) = 2t$, which is a time-varying coefficient. Independent right-censored data C_i was generated based in the exponential distribution to result 30 % of censoring. By (3.4), survival time X_i was generated. Also, observed time T_i and event indicator δ_i were derived by (3.1). Nine time knots were predetermined to be evenly spaced using the 10 percentiles of event time so that the same number of events was observed between each time knots using R package cox.spline. These 9 time knots were also applied to estimate pseudo-observations. The pseudo-observations, the estimated survival function based on Gray's time-varying model, and the estimated survival function based on Cox PH model were calculated for all individuals at 9 time knots.

True survival function was also obtained by direct input with $\theta = 1$, $\lambda = 1$, and $\beta(t)' = (2t_1, 2t_2, 2t_3, 2t_4, 2t_5, 2t_6, 2t_7, 2t_8, 2t_9)$, where t_k ($k = 1, \ldots, 9$) denoted predetermined time knots. In addition, the average of pseudo-observations at each time knot was calculated, and the estimate of survival function under the Kaplan-Meier estimator was derived by R package survfit.

3.2.2 The results

The true survival function, the average of pseudo-observations for survival function, the estimated survival function based on Kaplan-Meier estimator, the estimated survival function based on Gray's model, and the estimated survival function based on Cox model at 9 time knots (.070, .129, .187, .258, .310, .381, .447, .528, .651) for $Z = 0$ and $Z = 1$ are presented in Table 1 and Table 2. The results in Table 1 and Table 2 are plotted in Fogure 3, and Figure 4. The results support that the average of the pseudo-observations for the survival function computed at each time point, $\frac{1}{n}\sum_1^n S_i(t)$, is close to the estimated survival function based on Kaplan-Meier estimator, $\hat{S}_{KM}(t)$, and the true survival function, $S(t)$, [84]. The estimated survival probability based on Gray's time-varying coefficients model, $\hat{S}_{Gray}(t)$, is quite close to the true survival function, $S(t)$, and those two values are slight different from the average of pseudo-observations and the estimated survival function based on Kaplan-Meier estimator. The estimated survival function based on Cox PH model, $\hat{S}_{Cox}(t)$, is not close to the other estimates at most of the time points for $Z = 0$ and $Z = 1$. Based on Table 1 and Table 2, we can conclude that the pseudo-observations, $S_i(t)$, represents the true survival function well; therefore it can replace incomplete data. Also, Gray's model shows a good of fit for the model in violating of proportional hazards assumption.

Table 1: Simulation I Comparing the true survival function to the estimated average of pseudo-observations, the estimated survival based on Kaplan-Meier estimator, Gray's model, and Cox model

	Time Knots	**.070**	**.129**	**.187**	**.258**	**.310**	**.381**	**.447**	**.528**	**.651**
z=1	$S(t)$ **True value**	**.927**	**.863**	**.797**	**.713**	**.650**	**.565**	**.486**	**.392**	**.262**
	$\hat{S}_{KM}(t)$ **Kaplan-Meier**	**.916**	**.870**	**.788**	**.722**	**.639**	**.561**	**.441**	**.353**	**.224**
	$\frac{1}{n}\sum_1^n S_i(t)$ **Average of Pseudo-observations**	**.916**	**.870**	**.788**	**.722**	**.641**	**.566**	**.442**	**.356**	**.227**
	$\hat{S}_{Gray}(t)$ **Gray's time-varying model**	**.933**	**.868**	**.797**	**.723**	**.645**	**.561**	**.467**	**.364**	**.228**
	$\hat{S}_{Cox}(t)$ **Cox PH model**	**.944**	**.884**	**.819**	**.748**	**.668**	**.581**	**.481**	**.364**	**.214**

Table 2: Simulation I Comparing the true survival function to the estimated average of pseudo-observations, the estimated survival based on Kaplan-Meier estimator, Gray's model, and Cox model

	Time Knots	**.070**	**.129**	**.187**	**.258**	**.310**	**.381**	**.447**	**.528**	**.651**
z=2	$S(t)$ **True value**	**.922**	**.845**	**.756**	**.637**	**.540**	**.408**	**.289**	**.163**	**.044**
	$\hat{S}_{KM}(t)$ **Kaplan-Meier**	**.933**	**.827**	**.740**	**.638**	**.535**	**.415**	**.328**	**.187**	**.065**
	$\frac{1}{n}\sum_1^n S_i(t)$ **Average of Pseudo-observations**	**.933**	**.827**	**.745**	**.639**	**.536**	**.416**	**.329**	**.190**	**.066**
	$\hat{S}_{Gray}(t)$ **Gray's time-varying model**	**.918**	**.830**	**.738**	**.638**	**.531**	**.421**	**.307**	**.185**	**.064**
	$\hat{S}_{Cox}(t)$ **Cox PH model**	**.909**	**.816**	**.719**	**.618**	**.514**	**.407**	**.299**	**.188**	**.078**

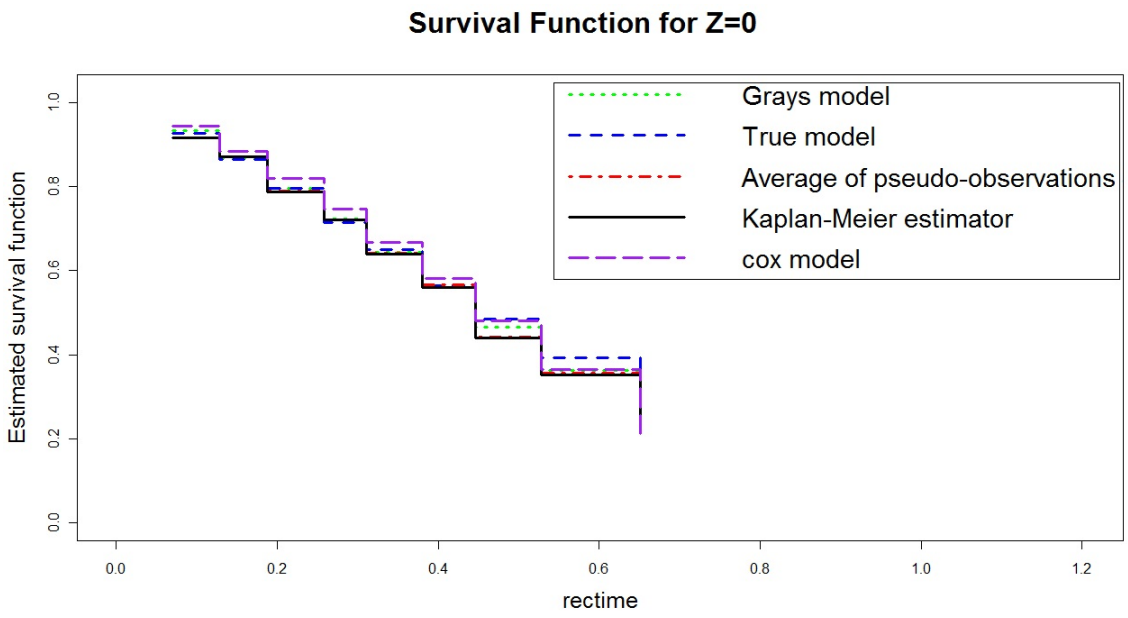

Figure 3: Simulation I Plot of survival function vs. time at $Z = 0$: True-value, Average of Pseudo-observation, Kaplan-Meier estimator, Gray's time-varying coefficients model, and Cox PH model

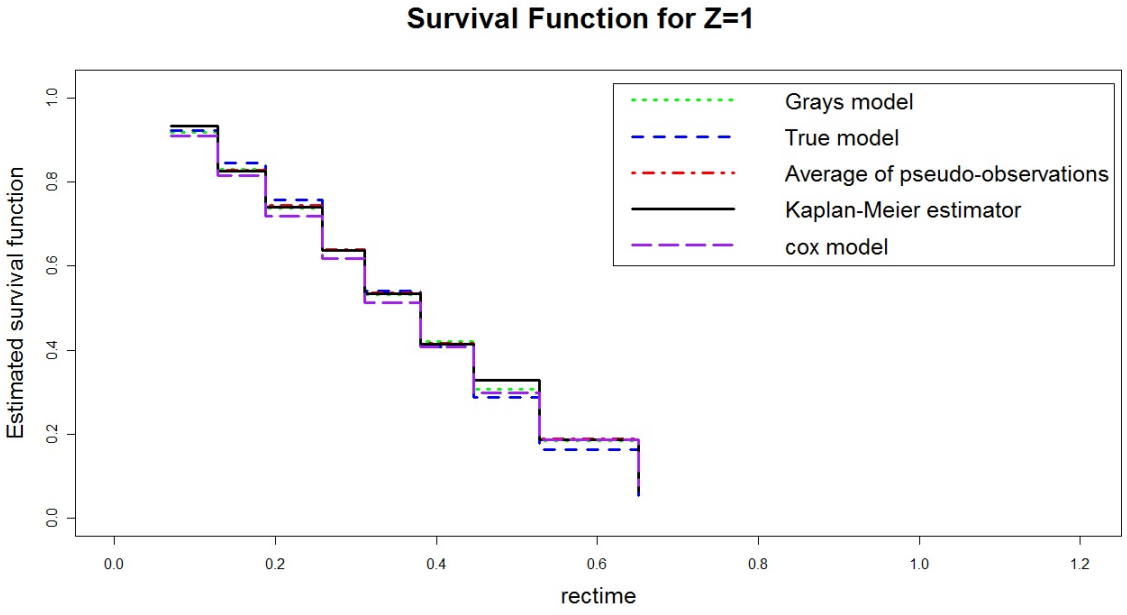

Figure 4: Simulation I Plot of survival function vs. time at $Z = 1$: True-value, Average of Pseudo-observation, Kaplan-Meier estimator, Gray's time-varying coefficients model, and Cox PH model

3.3 SIMULATION II: GOODNESS-OF-FIT TEST FOR GRAY'S PIECEWISE-CONSTANT TIME-VARYING COEFFICIENTS MODEL AND COX PH MODEL USING PSEUDO-OBSERVATIONS

3.3.1 Generate Data Set and Overview Methods

The goals of simulation II are to compare the estimated survival functions obtained from the Gray's time-varying coefficients model and from the Cox PH model to the true value, and to assess goodness-of-fit via the residual plots at each of the 9 time point. The risk factor Z for the i^{th} subject follows normal distribution, $Z \sim$ N(5,1), which is time-varying coefficient covariate. Let $h_0(t) = \lambda = 1$, which is a constant baseline hazard function, and $\theta = 2$ and $b(t) = 2t$, which is time-varying coefficient. Independent right-censored data C_i was generated based on the exponential distribution to result 0 %, 8.4 %, 37.8 %, and 71 % of censoring. Survival time X_i was generated following (3.4). Also, observed time T_i and event indicator δ_i were derived by following setting:

$T_i = min(X_i, C_i)$, the observation time on study for i^{th} subject

δ_i: the event indicator for i^{th} subject: 1 if an individual has experienced the event, 0 if an individual is censored

Nine time knots were predetermined to be evenly spaced using the 10 percentiles of event times so that the same number of events was observed between each time knot by R package cox.spline. These 9 time knots were also applied to estimate pseudo-observations. The estimated survival function based on Gray's time-varying coefficients models and based on Cox models, and true survival function for $Z = z$ were calculated. The true survival function was obtained by direct input with $\theta = 1$, $\lambda = 1$, and $\beta(t)' = (2t_1, 2t_2, 2t_3, 2t_4, 2t_5, 2t_6, 2t_7, 2t_8, 2t_9)$, where t_k $(k = 1, \ldots, 9)$ denoted predetermined time knots. These three survival functions were plotted with respect to time, and compared to see how close each other.

Also, the pseudo residuals using pseudo-observations for Gray's time-varying coefficients models and Cox PH models and true-Gray residual using true survival function for Gray's time-varying coefficiets models were calculated.

To assess goodness-of-fit test using residuals as a graphical tool, residual plots along with smoothed average plots were conducted against the estimated survival function based on fitted model at each 9 time knot.

3.3.2 Simulation Study II-1 No Censoring

Simulation II-1 was performed using survival data generated with $\theta = 2$, $\lambda = 1$, and no censoring. The results are presented in Table 3 and Figure 5 trough Figure 10. Test statistic for proportional hazards assumption is 12.64 ($p < .0001$). This test statistic shows that in true model, the effects of the covariate are varying over time.

The true regression coefficient, $\beta(t)$, the estimated regression coefficient based on Gray's time-varying coefficients model, $\hat{\beta}(t)$, and the estimated regression coefficient based on Cox PH model, $\hat{\beta}$, at each 9 time knot are presented in Table 3. While the estimated regression coefficient based on Cox's PH model is constant over time, .435 ($p < .001$), the estimated regression coefficient based on Gray's time-varying coefficients model (.150, .229, .324, .427, .493, .526, .493, .474, .471), and true regression coefficient (.140, .236, .306, .364, .418, .474, .532, .604, .668) are changing over time. The estimated regression coefficient based on Gray's model is close to true regression coefficient except for early time point. The results of Table 3 are plotted against time in Figure 5 and Figure 6.

The plot of log-hazard-ratio for the covariate Z with respect to time is presented in Figure 5. By (2.14) , Figure 5 represents the trend of $\beta(t)$ over time. The result shows that the effect of this covariate is changing over time, i.e., time-varying coefficient. Plots of the estimated regression coefficient based on Gray's time-varying coefficients model and based on Cox PH model to the true regression coefficient were compared in Figure 6.

The result shows that the estimated regression coefficient based on Gray's model is close to the true regression coefficient. Also, the estimated survival functions were calculated and plotted based on Gray's time-varying coefficients model and based on Cox PH model at $(Z = 2, 4, 6, 8)$ (Figure 7). Comparing the estimated survival function based on Gray's model to the true survival function, Gray's time-varying coefficients model is found to be a good fit in estimating survival function.

If there is no censoring (complete data), Gray's model shows the best fit when the covariate Z has near mean value of Z. Compared to true survival function, the plot from Gray's model show departure at the end of time when covariate Z has small value, and departure at the beginning of time when the covariate Z has large value .

In this simulation study, 3 types of residuals were calculated using pseudo-observations and true survival function. Pseudo residual for Gray's time-varying coefficients model is defined by

$$PseudoGray\ residual = S_i(t) - \hat{S}_{Gray}(t|Z_i)\ , \tag{3.5}$$

pseudo residual for Cox PH model is defined by

$$PseudoCox\ residual = S_i(t) - \hat{S}_{Cox}(t|Z_i)\ , \tag{3.6}$$

and residual using true survival function_ estimated survival function based on Gray's model is defined by

$$TrueGray\ residual = S_{True}(t|Z_i) - \hat{S}_{Gray}(t|Z_i)\ . \tag{3.7}$$

The residual plots along with smoothed average against the estimated survival rate based on fitted model at each 9 time point are presented in Figure 8, Figure 9 and Figure 10.

In Figure 8, the pseudo residuals are plotted against the estimated survival rate based on Gray's model at each 9 time point to assess goodness-of-fit for Gray's time-varying coefficicents model using the pseudo-observation. Since the pseudo residual plots for Gray's model are constant near zero at all time points, the conclusion is that Gray's time-varying coefficients model shows a good fit in estimating survival in case of violating of proportional hazards assumption.

In Figure 9, the true-Gray residuals is plotted against the estimated survival rate based on Gray's time-varying coefficients model at each 9 time point to assess goodness-of-fit for Gray's model, $\hat{S}_{Gray}(t|Z_i)$, using the true survival function, $S_{True}(t|Z_i)$. Since the residual plots are constant near zero, the conclusion is that Gray's time-varying coefficients model shows a good fit in estimating survival in case of violating of proportional hazards assumption.

Table 3: Simulation II-1 Compare true value of covariate effect $\beta(t)$ to estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and no censoring

Time Knots	**.070**	**.118**	**.153**	**.182**	**.209**	**.237**	**.266**	**.302**	**.334**
True $\beta(t)$: θ x Time (θ=2)	**.140**	**.236**	**.306**	**.364**	**.418**	**.474**	**.532**	**.604**	**.668**
$\hat{\beta}(t)$ based on Gray's model	**.150**	**.229**	**.324**	**.427**	**.493**	**.526**	**.493**	**.474**	**.471**
$\hat{\beta}$ based on Cox model	**.390**	**.390**	**.390**	**.390**	**.390**	**.390**	**.390**	**.390**	**.390**

In Figure 10, the pseudo-Cox residuals are plotted against the estimated survival rate based on Cox PH model using the pseudo-observation. Since the pseudo residual plots for Cox's model show departures and tendencies, the conclusion is that Cox PH model does not fit in estimating survival function in case of violating of proportional hazards assumption.

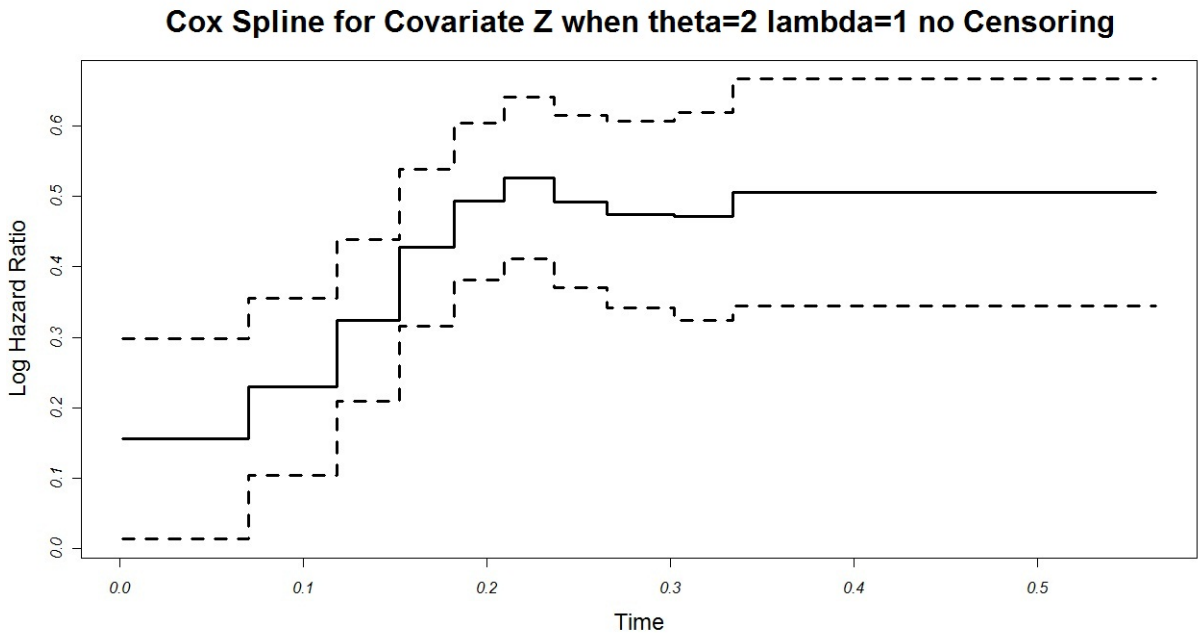

Figure 5: Simulation II-1 Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model (θ=2, λ=1, and no censoring)

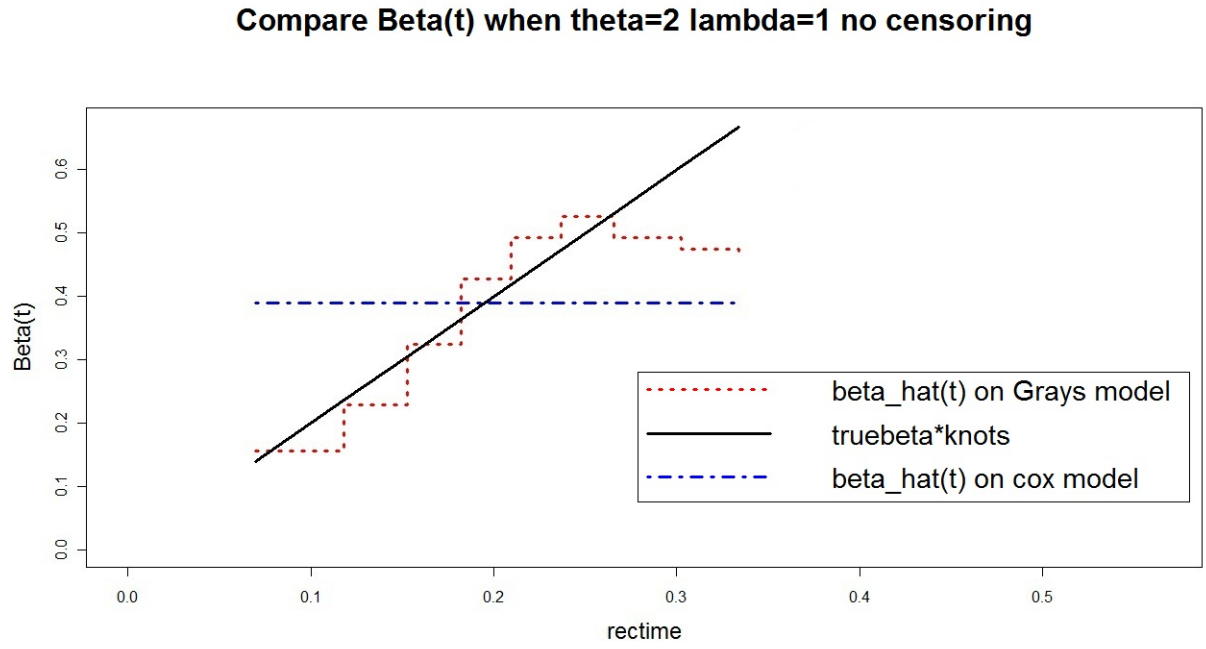

Figure 6: Simulation II-1 Plot of the true value of covariate effect $\beta(t)$, the estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model vs. time; true value of covariate effect $\beta(t) = 2t, \lambda = 1$, and no censoring

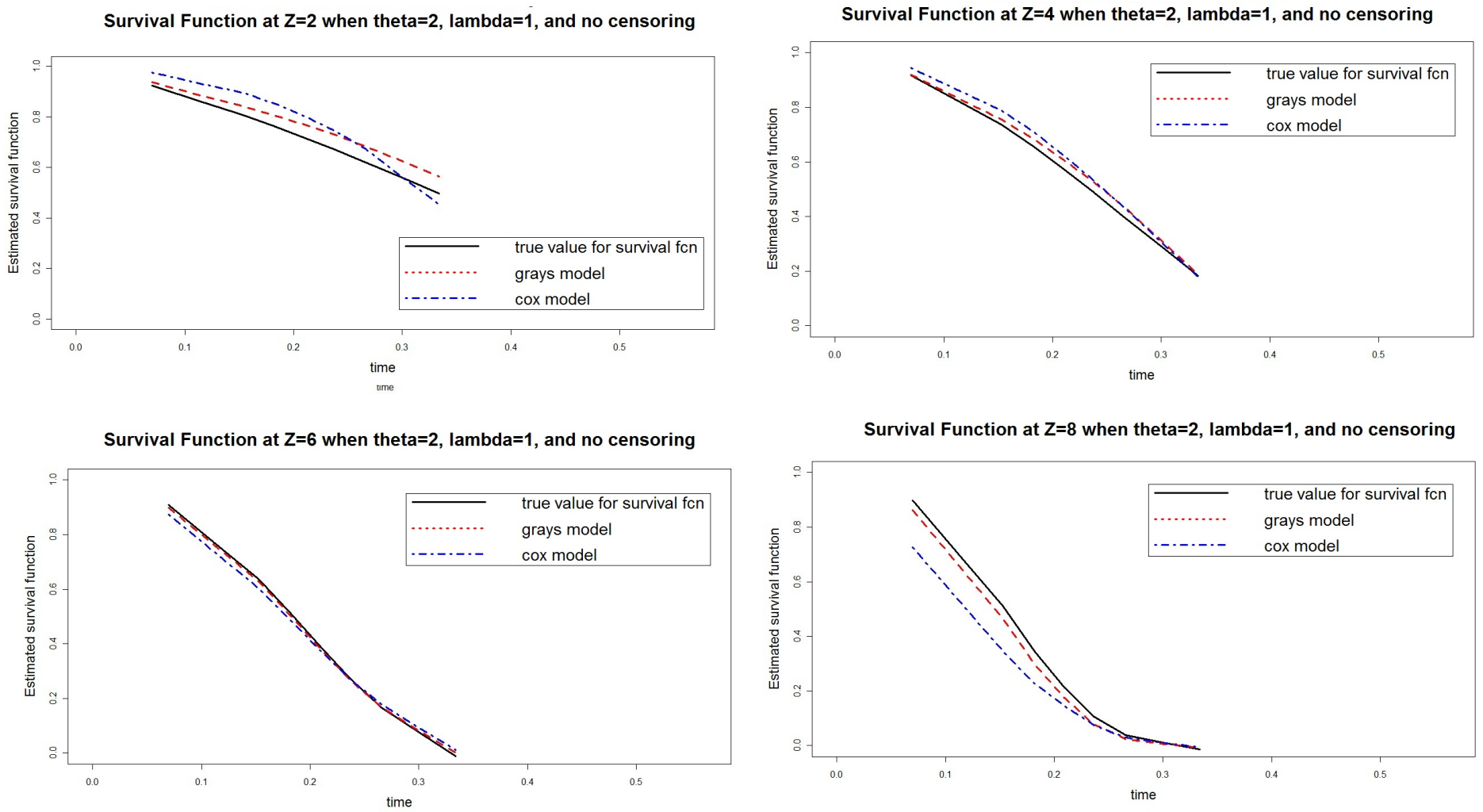

Figure 7: Simulation II-1 Plot of the estimated survival function based on Gray's time-varying coefficients and Cox PH model, and true survival function at (z=2,4,6, and 8): true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and no censoring

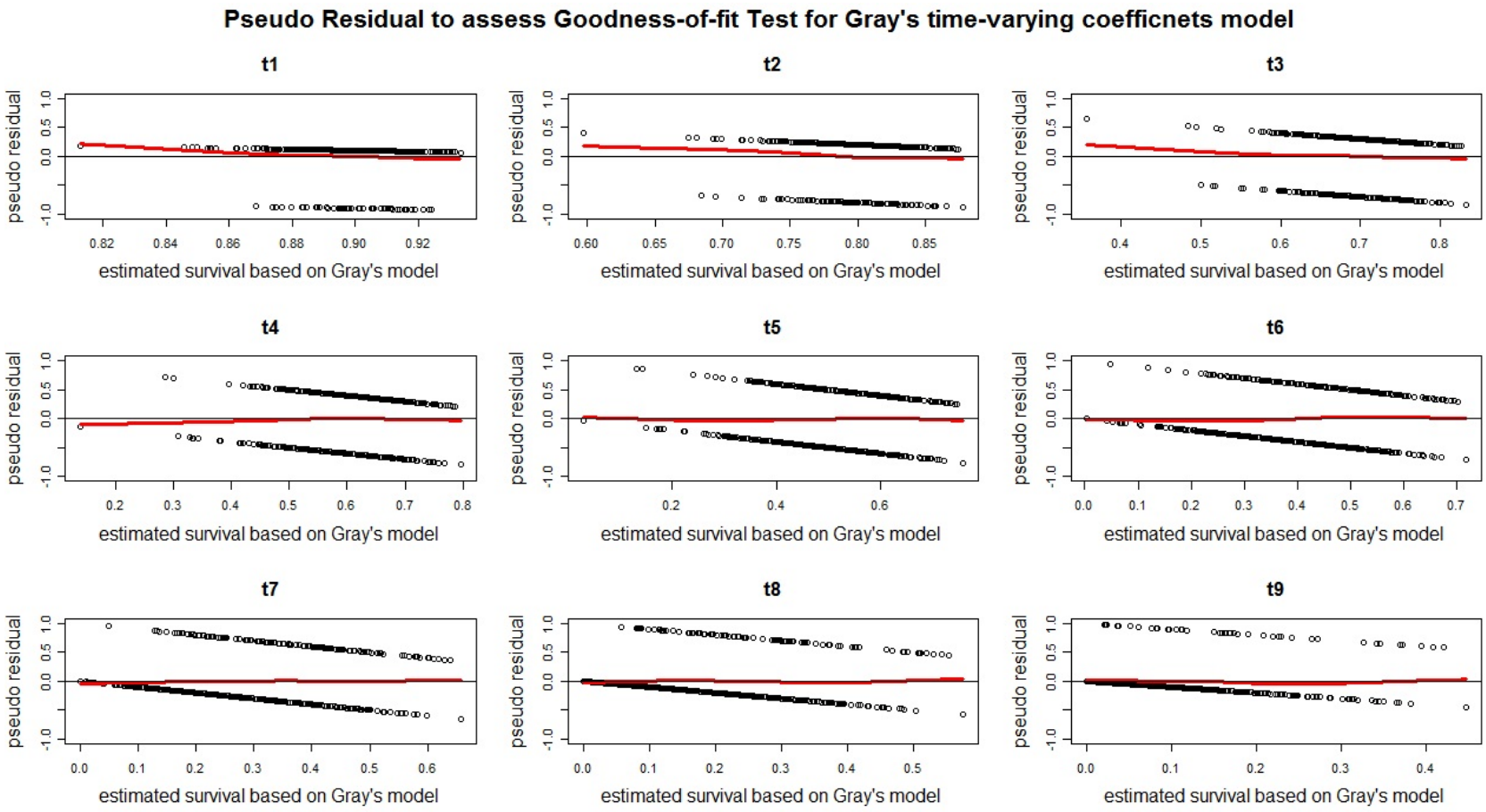

Figure 8: Simulation II-1 Pseudo residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and no censoring

$$PseudoGray\ residual = S_i(t) - \hat{S}_{Gray}(t|Z_i)$$

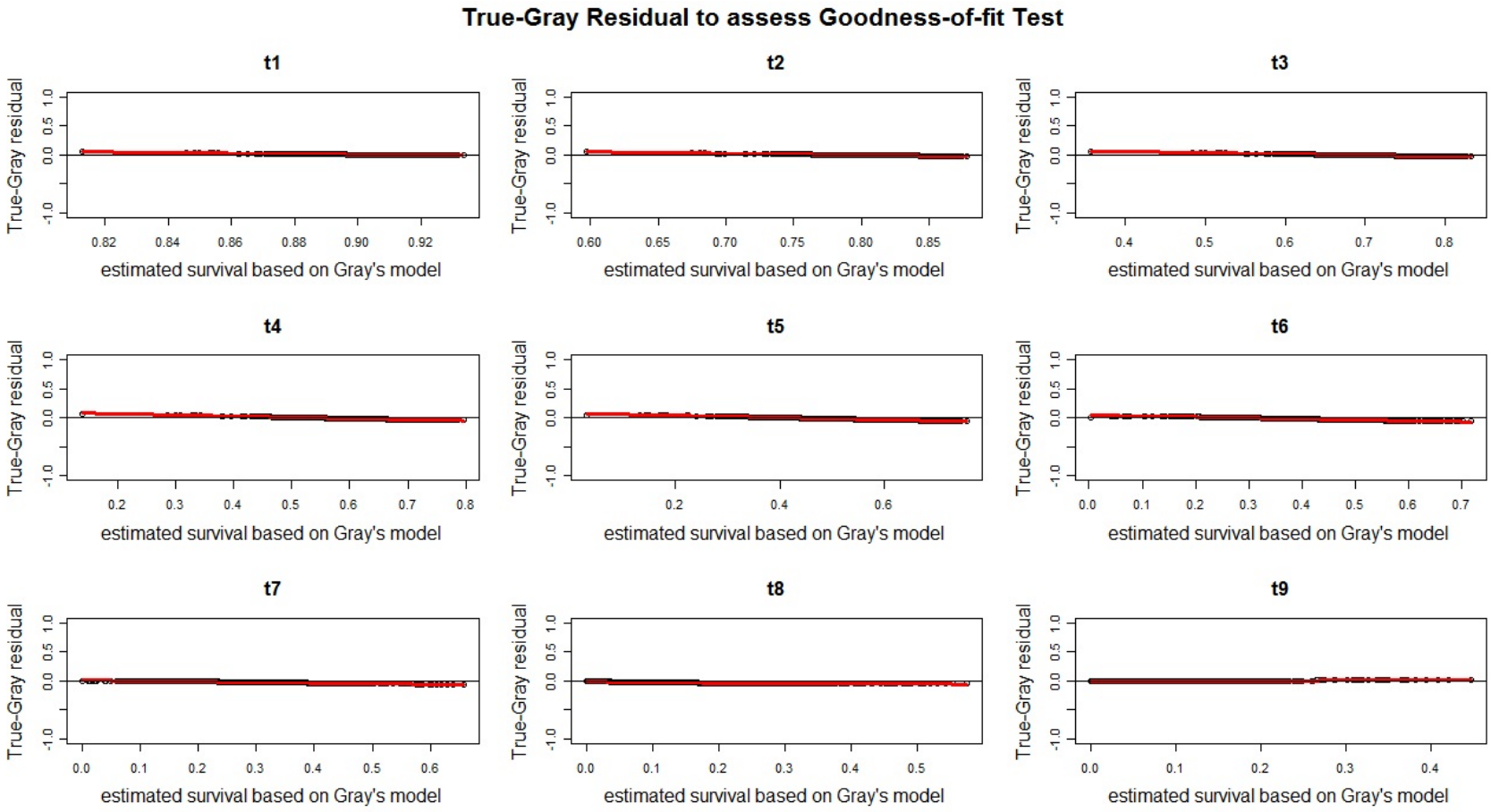

Figure 9: Simulation II-1 True-Gray residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and no censoring

$$TrueGray\ residual = S_{True}(t|Z_i) - \hat{S}_{Gray}(t|Z_i)$$

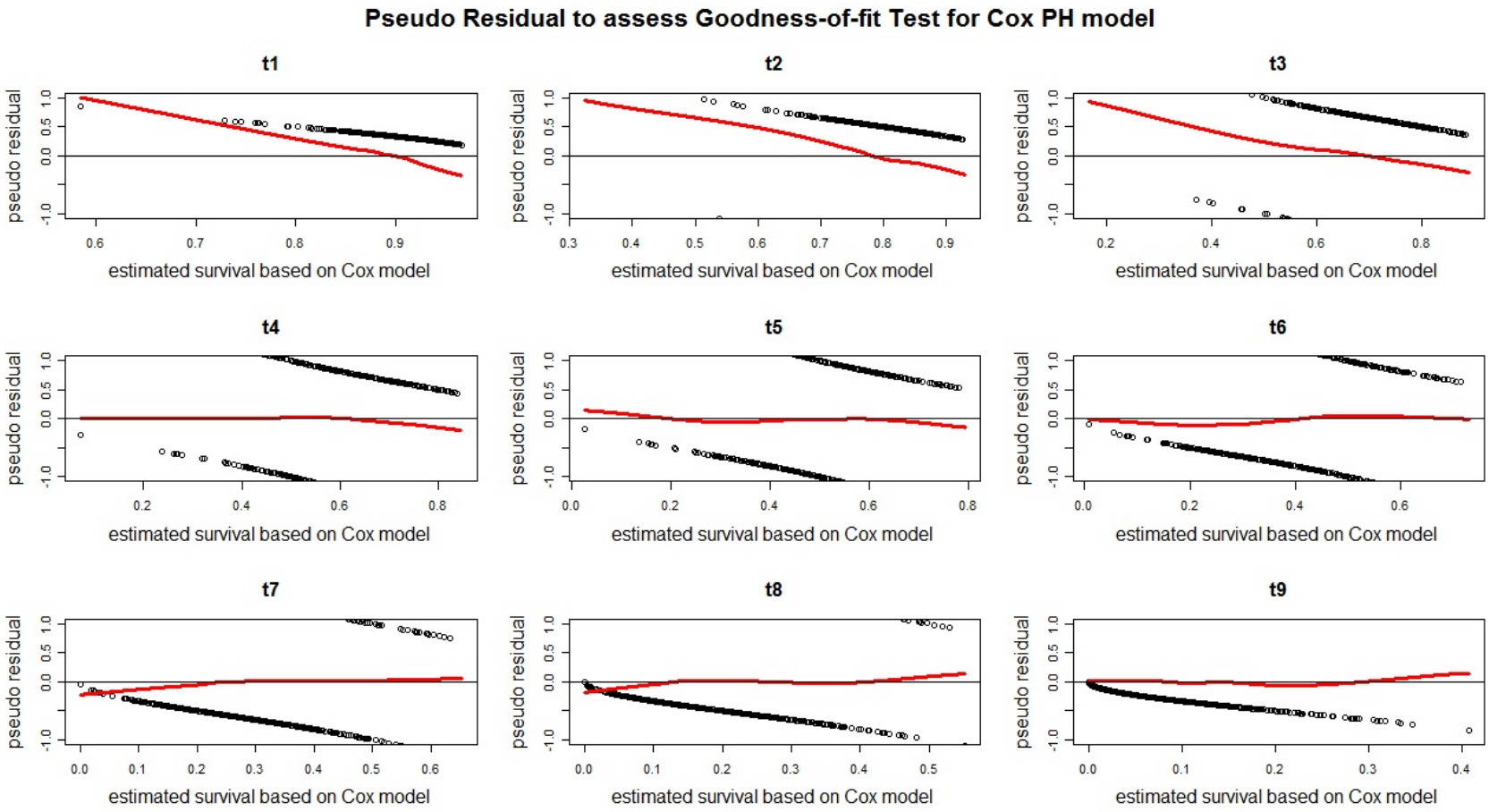

Figure 10: Simulation II-1 Pseudo residual vs. the estimated survival function based on Cox PH model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and no censoring

$$PseudoCox\ residual = \hat{S}_i(t) - \hat{S}_{Cox}(t|Z_i)$$

3.3.3 Simulation Study II-2 Light Censoring

Simulation II-2 was performed using survival data generated with $\theta = 2$, $\lambda = 1$, and light censoring (8.4 %). The results are presented in Table 4 and Figure 11 trough Figure 16. The value of the test statistic for proportional hazards assumption is 11.45 (p =.0002). This test statistic states that in true model, the effects of the covariate are varying over time.

The true regression coefficient, the estimated regression coefficient based on Gray's time-varying coefficients model, and the estimated regression coefficient based on Cox PH model at each 9 time knot are presented in Table 4. While the estimated regression coefficient based on Cox's model are constant over time, .037 ($p < .001$), the estimated regression coefficient based on Gray's model (.181, .210, .267, .306, .364, .413, .476, .535, .604), and true regression coefficient (.110, .214, .290, .346, .392, .448, .488, .554, .634) are changing over time. The estimated regression coefficients based on Gray's model is close to true regression coefficient except for early time points. The results of Table 4 were plotted against time in Figure 11 and Figure 12. The estimated survival function based on Gray's time-varying coefficients model and based on Cox PH model were calculated and plotted against time at $Z = (2, 4, 6, 8)$ (Figure 13). Finally, pseudo residual plots and true-Gray residual plots along with smoothed average against the estimated survival rate based on fitted model at each 9 time point were conducted and presented in Figure 14, Figure 15, and Figure 16 to assess the goodness-of-fit for Gray's time-varying coefficients model and Cox PH model.

Table 4: Simulation II-2 Compare true value of covariate effect $\beta(t)$ to estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.4 % censoring

Time Knots	**.055**	**.107**	**.145**	**.173**	**.196**	**.224**	**.244**	**.277**	**.317**
True $\beta(t)$: θ x Time (θ=2)	**.110**	**.214**	**.290**	**.346**	**.392**	**.448**	**.488**	**.554**	**.634**
$\hat{\beta}(t)$ based on Gray's model	**.181**	**.210**	**.267**	**.306**	**.364**	**.413**	**.476**	**.535**	**.604**
$\hat{\beta}$ based on Cox model	**.386**	**.386**	**.386**	**.386**	**.386**	**.386**	**.386**	**.386**	**.386**

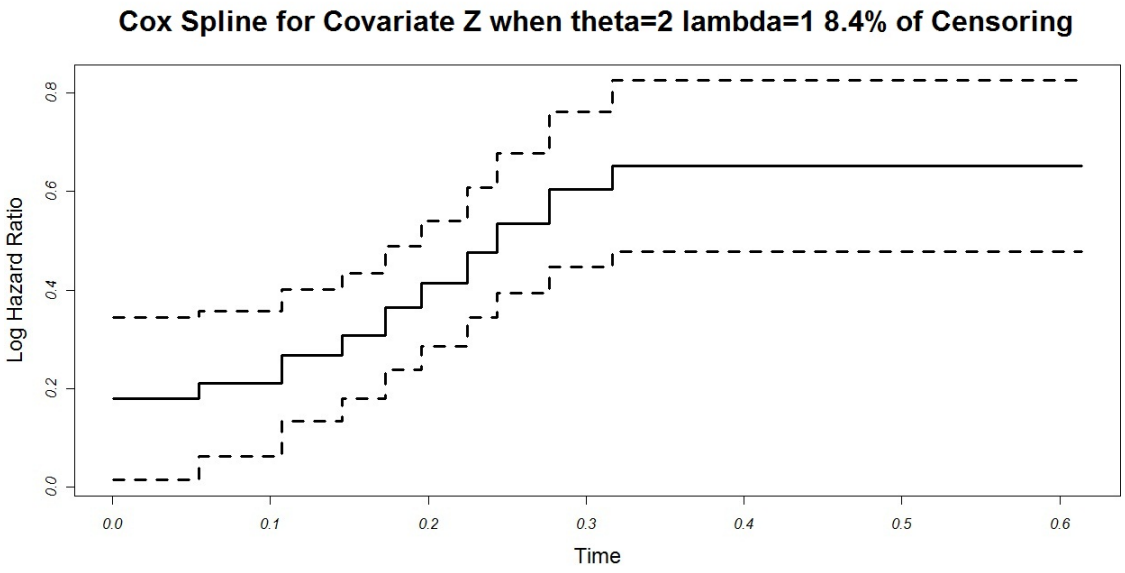

Figure 11: Simulation II-2 Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model (θ=2, λ=1, and 8.6 % censoring)

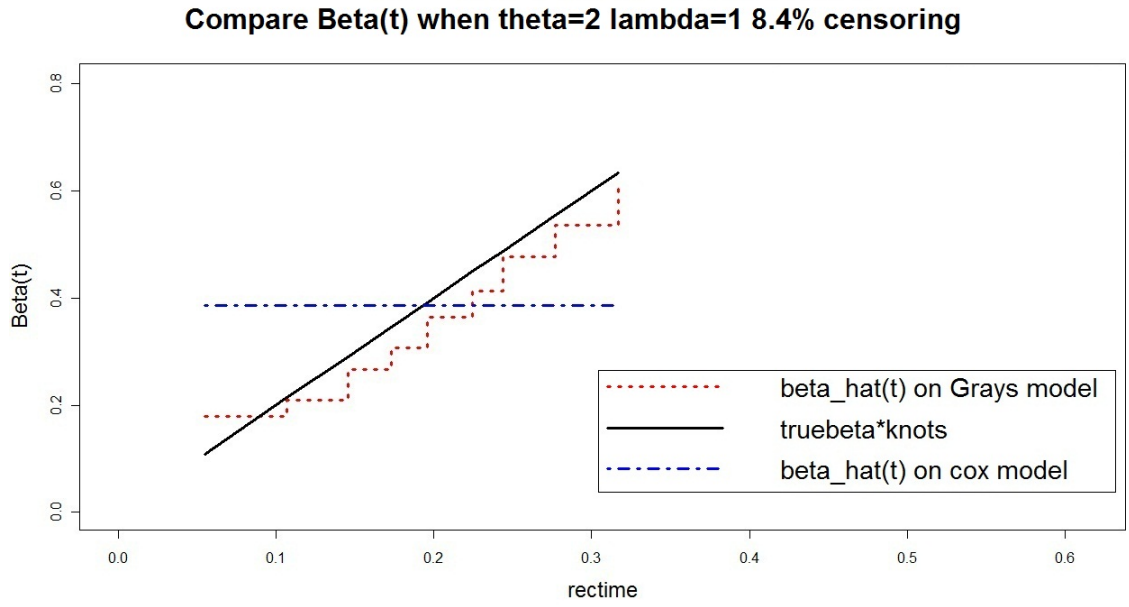

Figure 12: Simulation II-2 Plot of the true value of covariate effect $\beta(t)$, the estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model vs. time; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.6% censoring

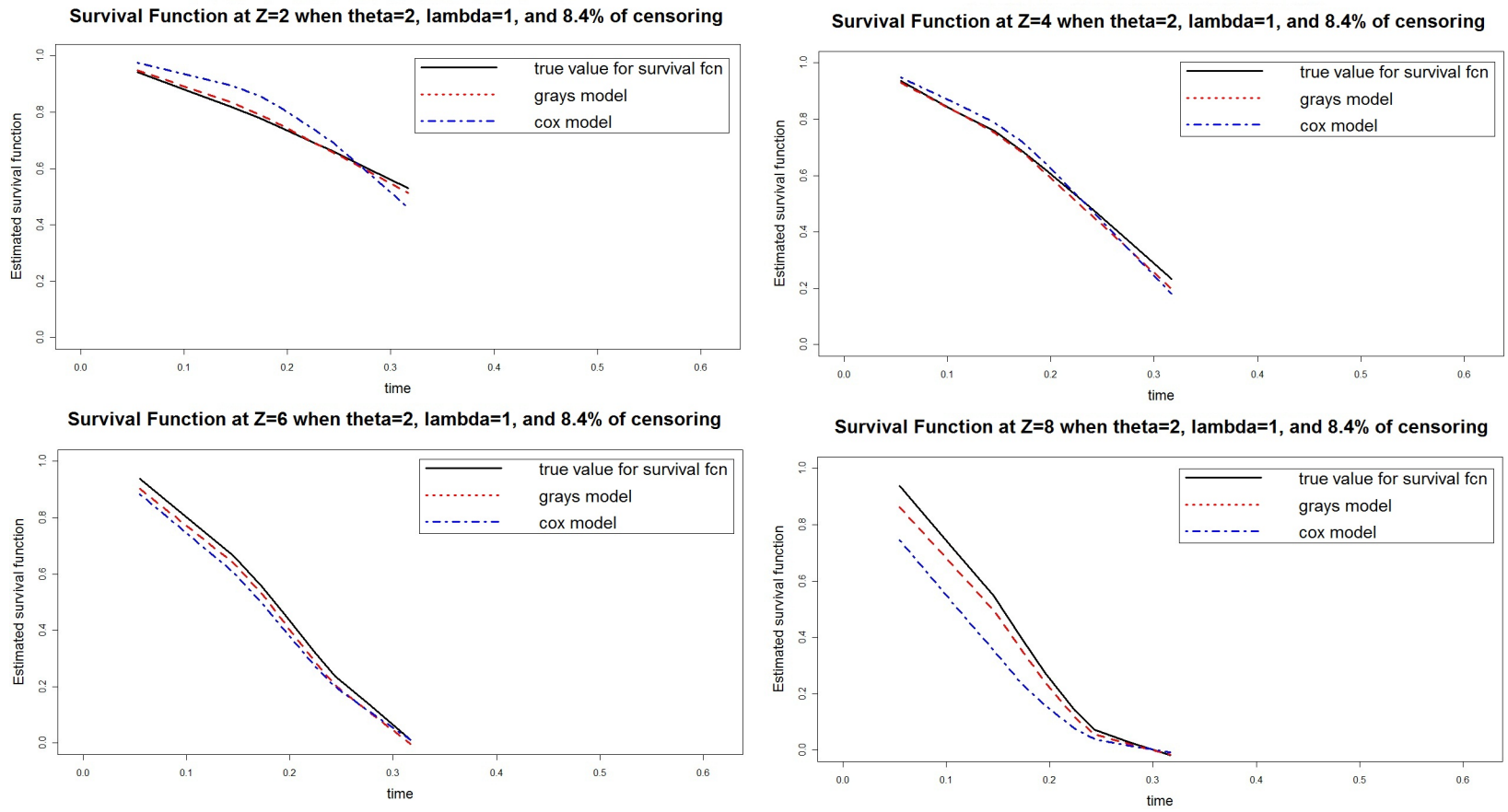

Figure 13: Simulation II-2 Plot of the estimated survival function based on Gray's time-varying coefficients and Cox PH model, and true survival function at (z=2,4,6, and 8): ; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.6 % censoring

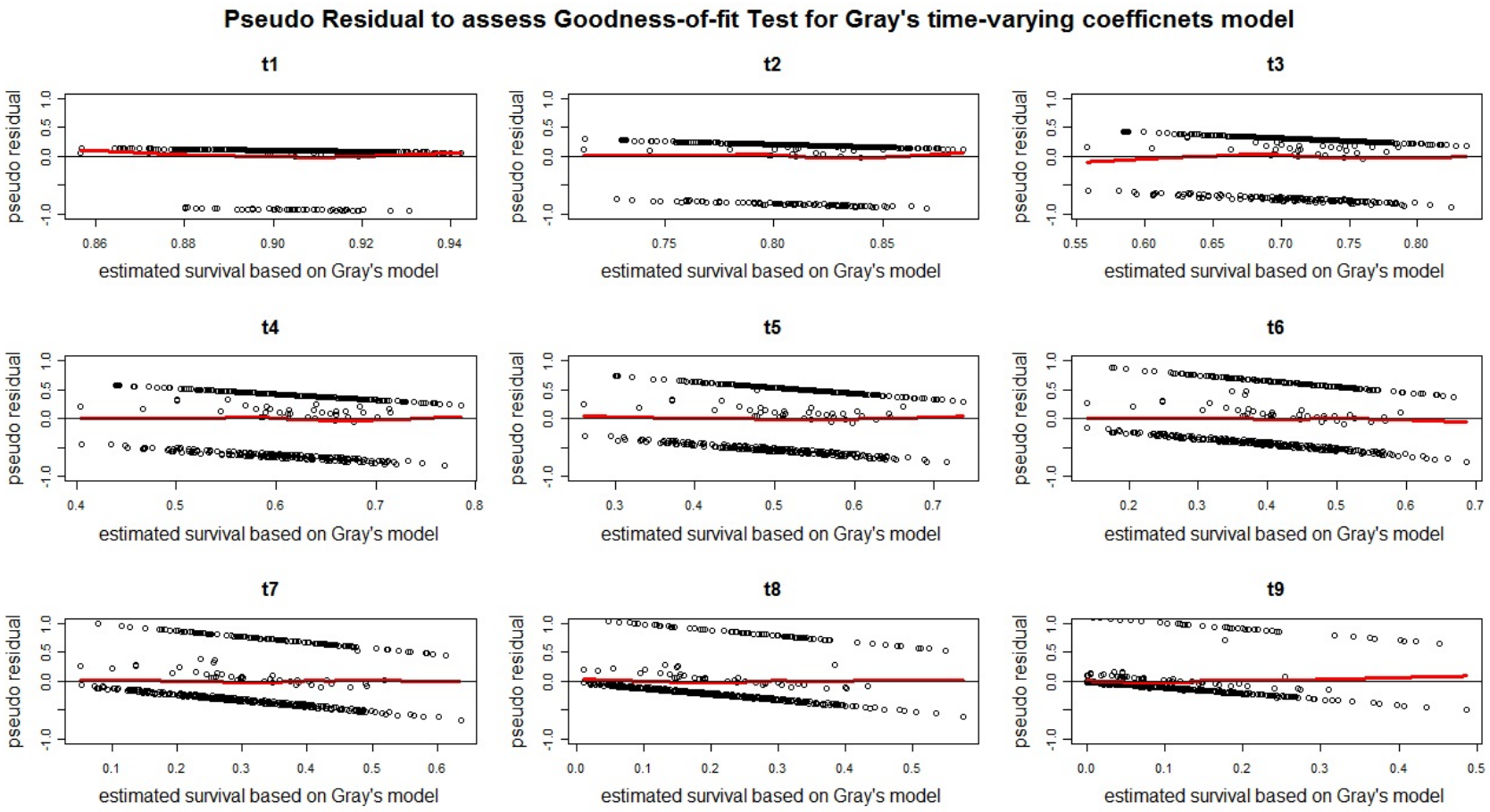

Figure 14: Simulation II-2 Pseudo residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.6 % censoring

$$PseudoGray\ residual = \hat{S}_i(t) - \hat{S}_{Gray}(t|Z_i)$$

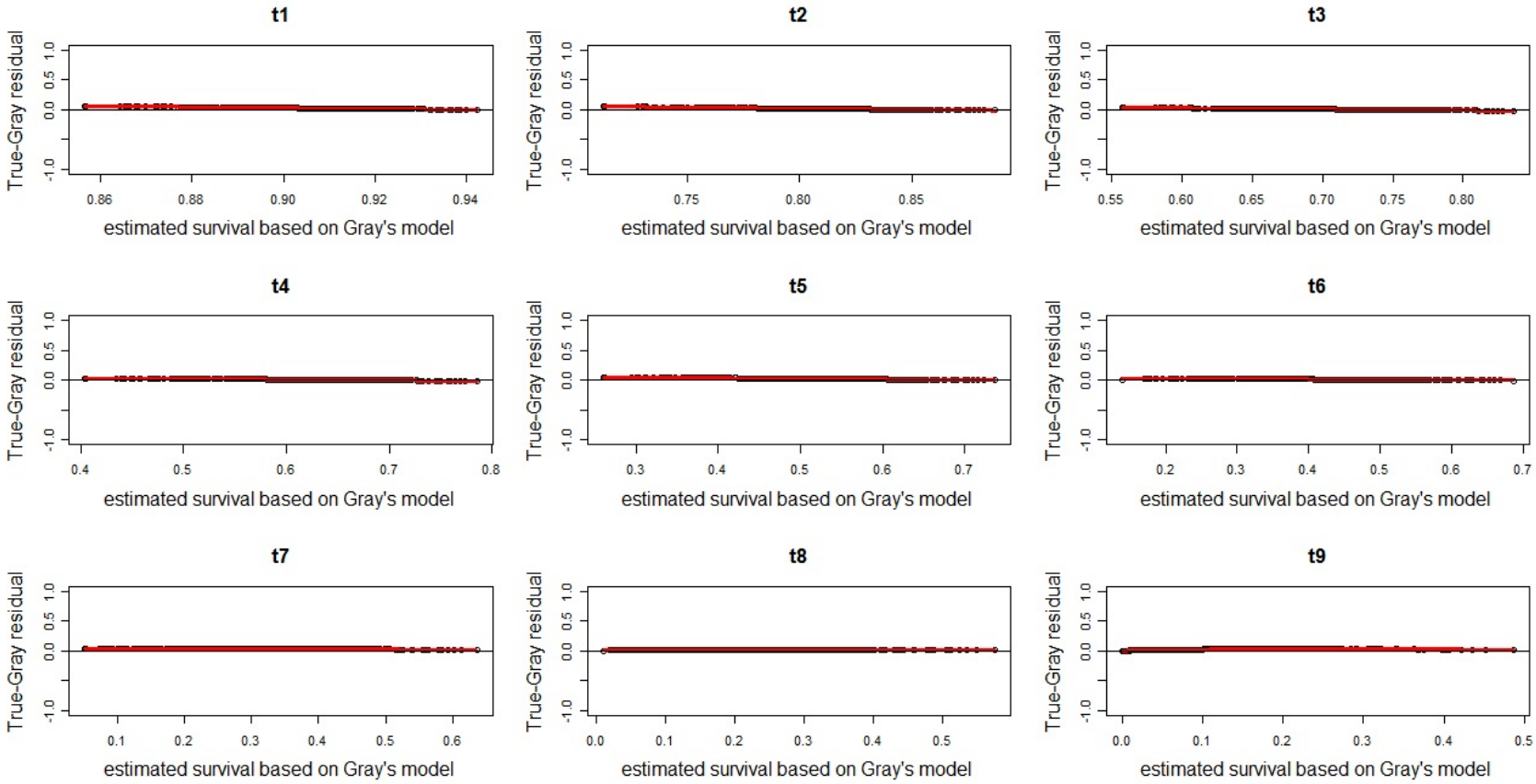

Figure 15: Simulation II-2 True-Gray residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.6 % censoring

$$TrueGray\ residual = S_{True}(t|Z_i) - \hat{S}_{Gray}(t|Z_i)$$

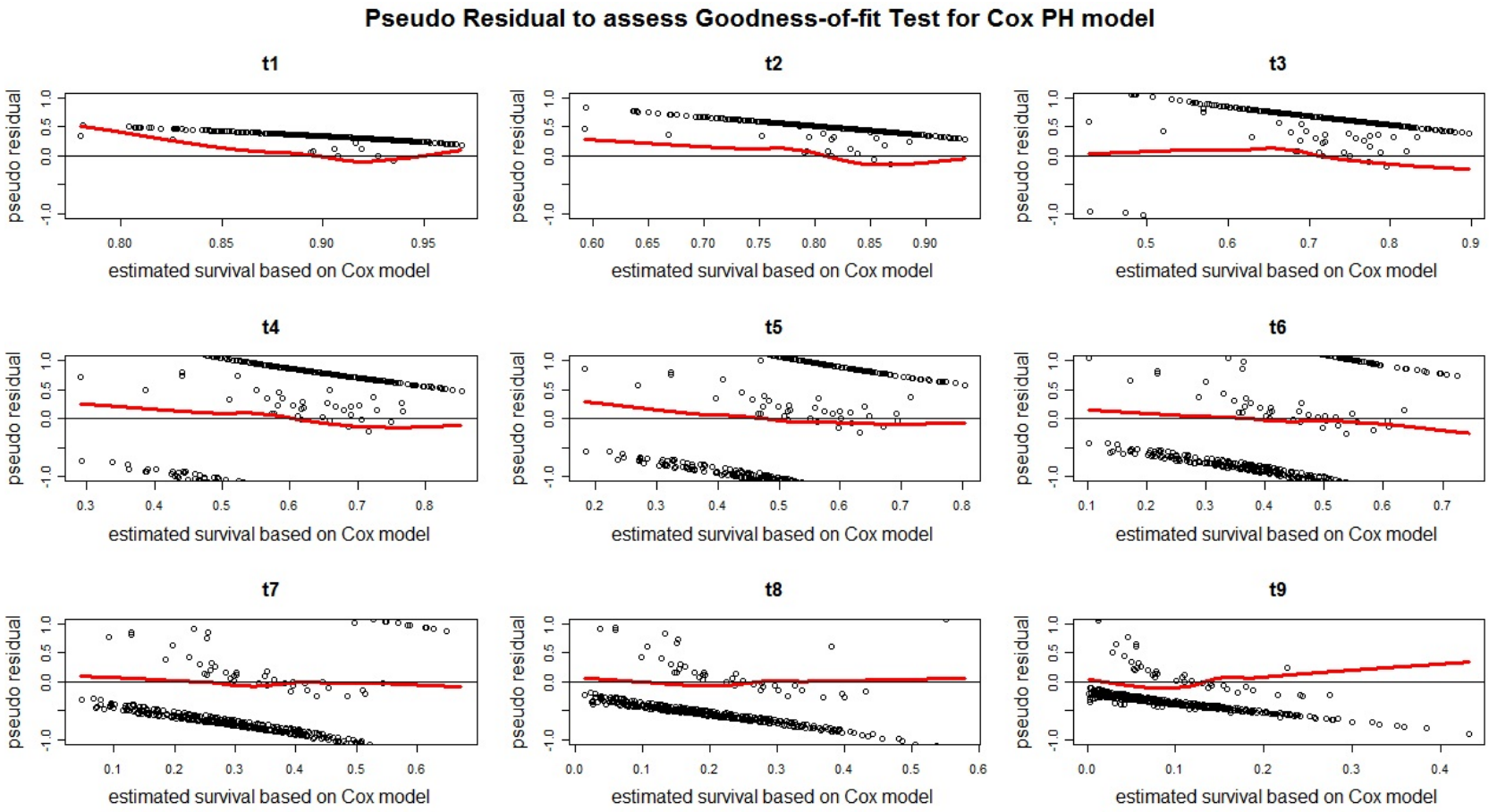

Figure 16: Simulation II-2 Pseudo residual vs. the estimated survival function based on Cox PH model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.6 % censoring

$$PseudoCox\ residual = \hat{S}_i(t) - \hat{S}_{Cox}(t|Z_i)$$

3.3.4 Simulation Study II-3 Mild Censoring

Simulation II-3 was performed using survival data generated with $\theta = 2$, $\lambda = 1$, and mild censoring (37.8 %). The results are presented in Table 5 and Figure 17 trough Figure 22. The value of the test statistic for proportional hazards assumption is 12.34 (p <.0001). This test statistic states that in true model, the effects of the covariate are varying over time.

The true regression coefficient, the estimated regression coefficient based on Gray's time-varying coefficients model, and the estimated regression coefficient based on Cox PH model at each 9 time knot are presented in Table 5. While the estimated regression coefficient based on Cox's model are constant over time, .325 (p<.001), the estimated regression coefficient based on Gray's model, (.063, .129, .230, .303, .358, .459, .538, .559, .538), and true regression coefficient, (.094, .154, .2.12, .292, .354, .410, .462, .522, .608), are changing over time. The estimated regression coefficients based on Gray's model is close to true regression coefficient except for early time point. The results of Table 5 were plotted in Figure 17 and Figure 18. The estimates of survival functions were calculated and plotted based on Gray's piecewise-constant time-varying model and Cox PH model at ($Z = 2, 4, 6, 8$) (Figure 19). Finally, pseudo residual plots and true-Gray residual plots along with smoothed average against the estimated survival rate based on fitted model at each 9 time point were conducted and presented in Figure 20, Figure 21, and Figure 22 to assess the goodness-of-fit for Gray's time-varying coefficients model and Cox PH model.

Table 5: Simulation II-3 Compare true value of covariate effect $\beta(t)$ to estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 37.8 % censoring

Time Knots	**.047**	**.077**	**.106**	**.146**	**.177**	**.205**	**.231**	**.261**	**.304**
True $\beta(t)$: θ x Time (θ=2)	**.094**	**.154**	**.2.12**	**.292**	**.354**	**.410**	**.462**	**.522**	**.608**
$\hat{\beta}(t)$ based on Gray's model	**.063**	**.129**	**.230**	**.303**	**.358**	**.459**	**.538**	**.559**	**.538**
$\hat{\beta}$ based on Cox model	**.353**	**.353**	**.353**	**.353**	**.353**	**.353**	**.353**	**.353**	**.353**

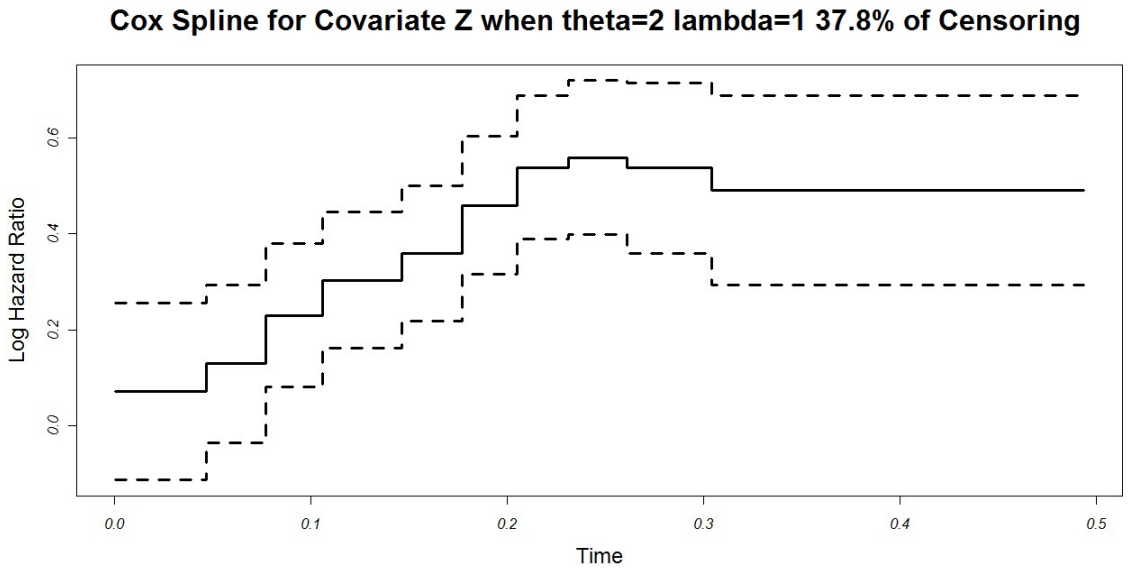

Figure 17: Simulation II-3 Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model (θ=2, λ=1, and 34.6% censoring)

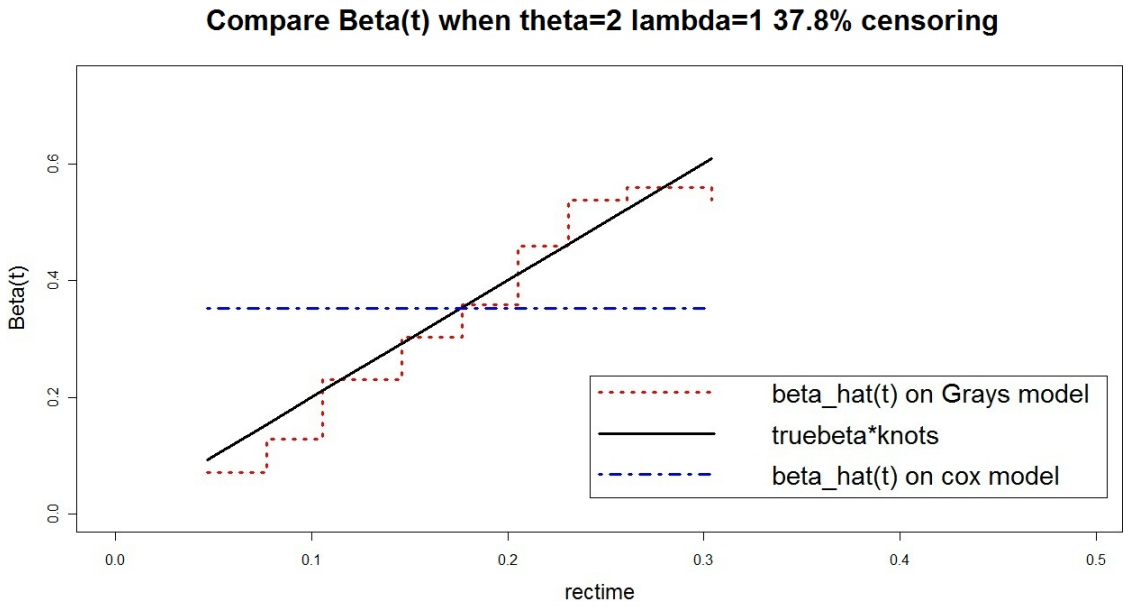

Figure 18: Simulation II-3 Plot of the true value of covariate effect $\beta(t)$, the estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model vs. time; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 8.6% censoring

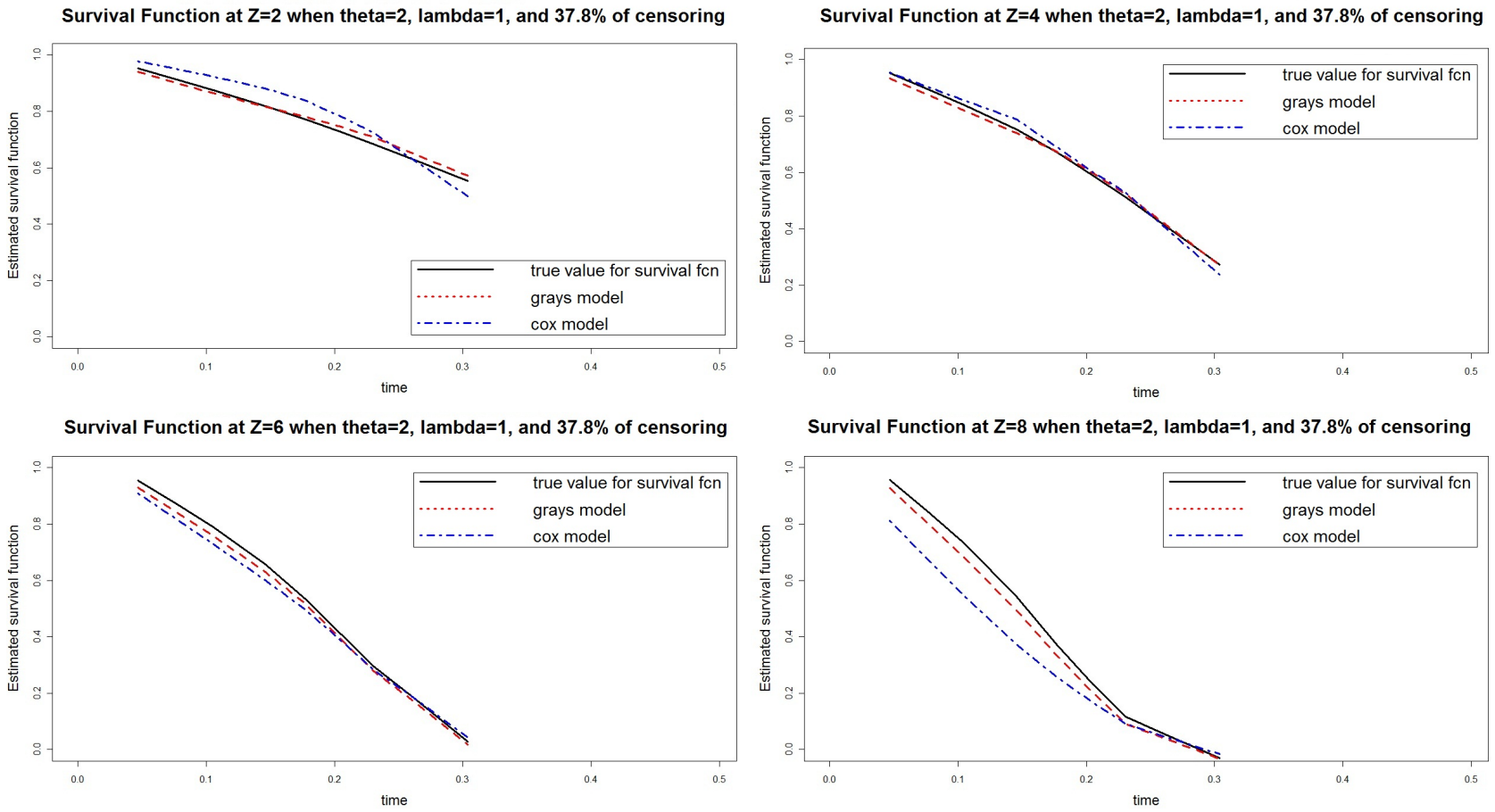

Figure 19: Simulation II-3 Plot of the estimated survival function based on Gray's time-varying coefficients and Cox PH model, and true survival function at (z=2,4,6, and 8): ; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 34.6 % censoring

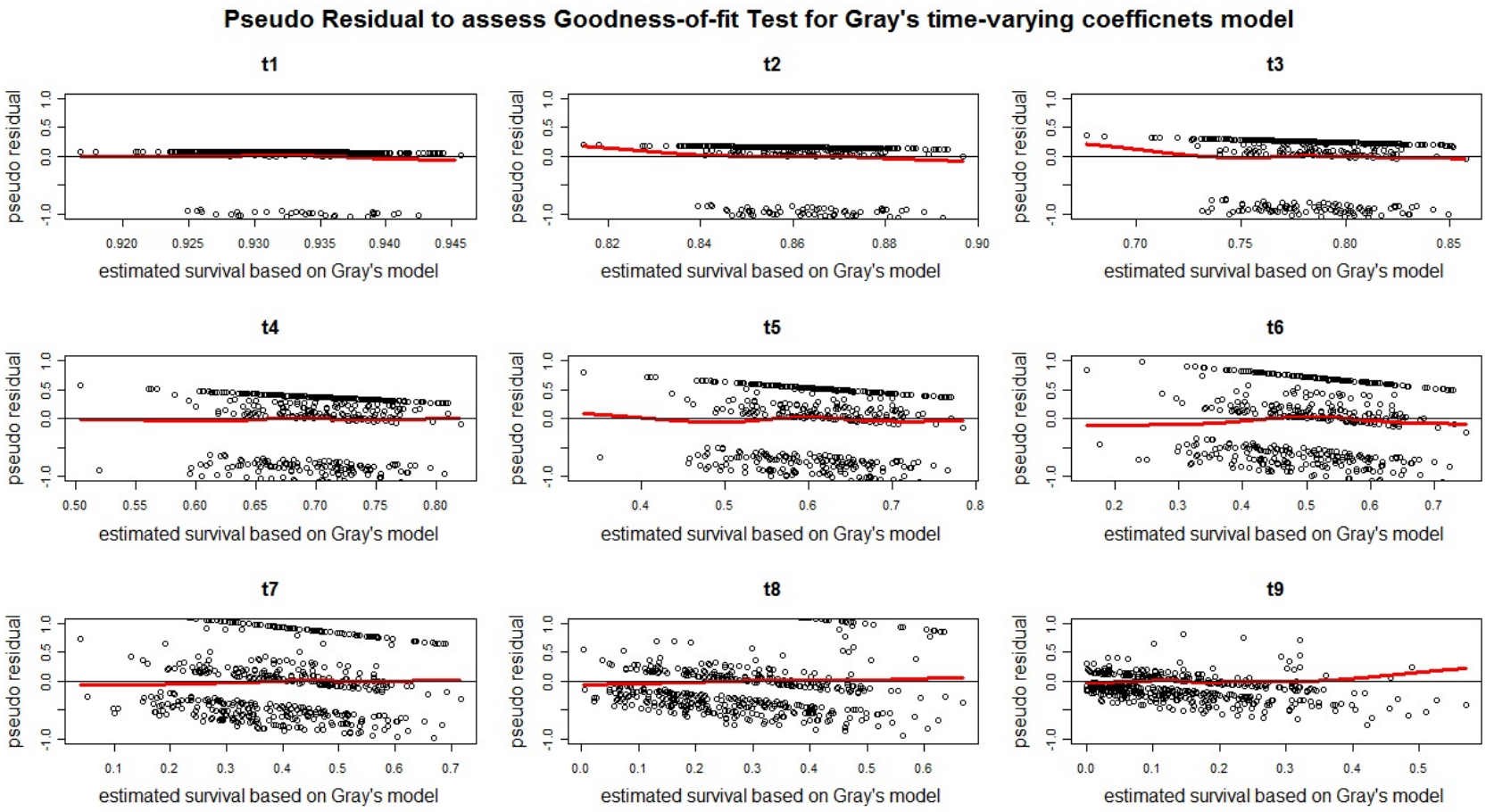

Figure 20: Simulation II-3 Pseudo residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 34.6 % censoring

$$PseudoGray\ residual = \hat{S}_i(t) - \hat{S}_{Gray}(t|Z_i)$$

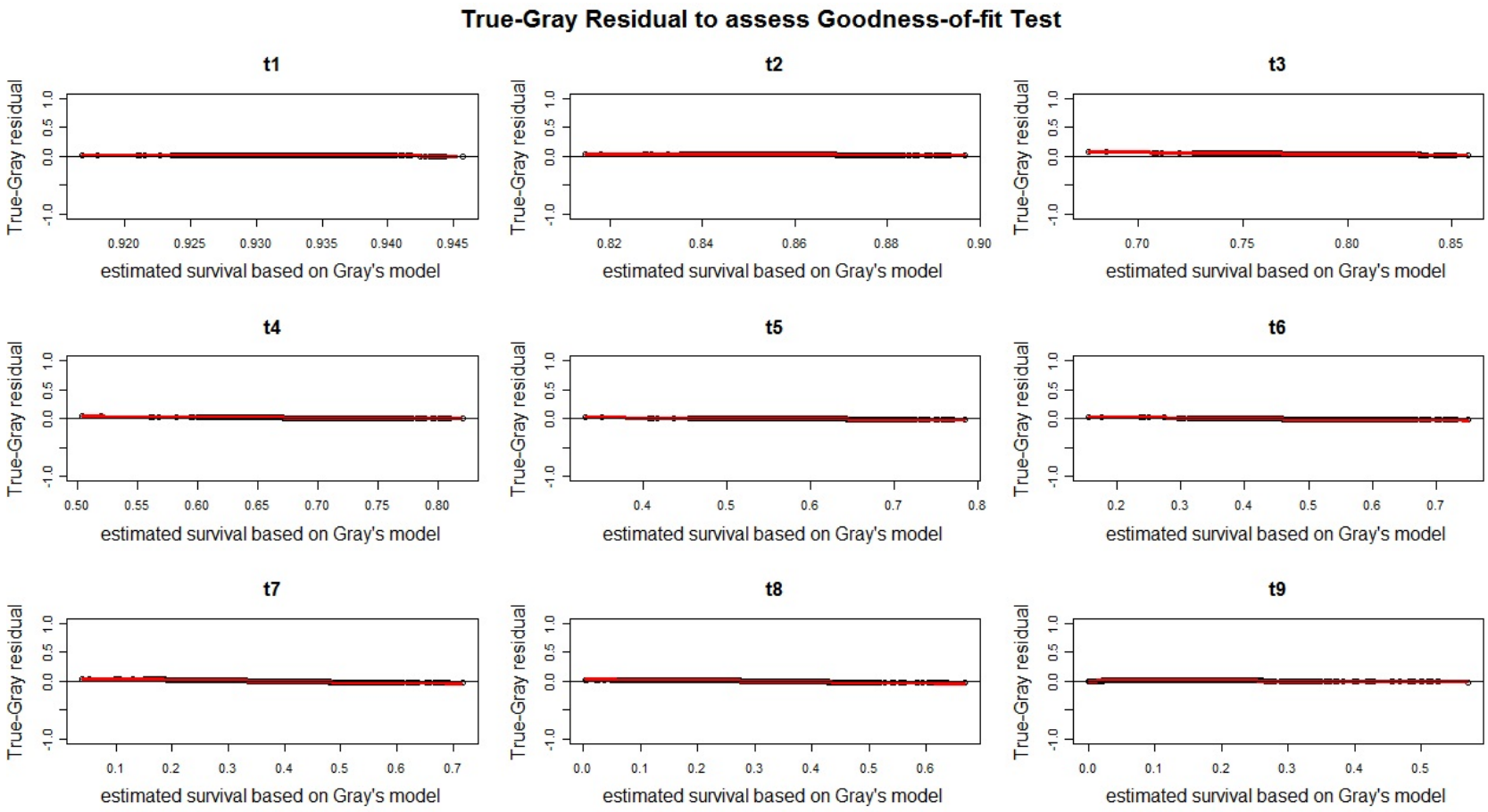

Figure 21: Simulation II-3 True-Gray residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 34.6 % censoring

$$TrueGray\ residual = S_{True}(t|Z_i) - \hat{S}_{Gray}(t|Z_i)$$

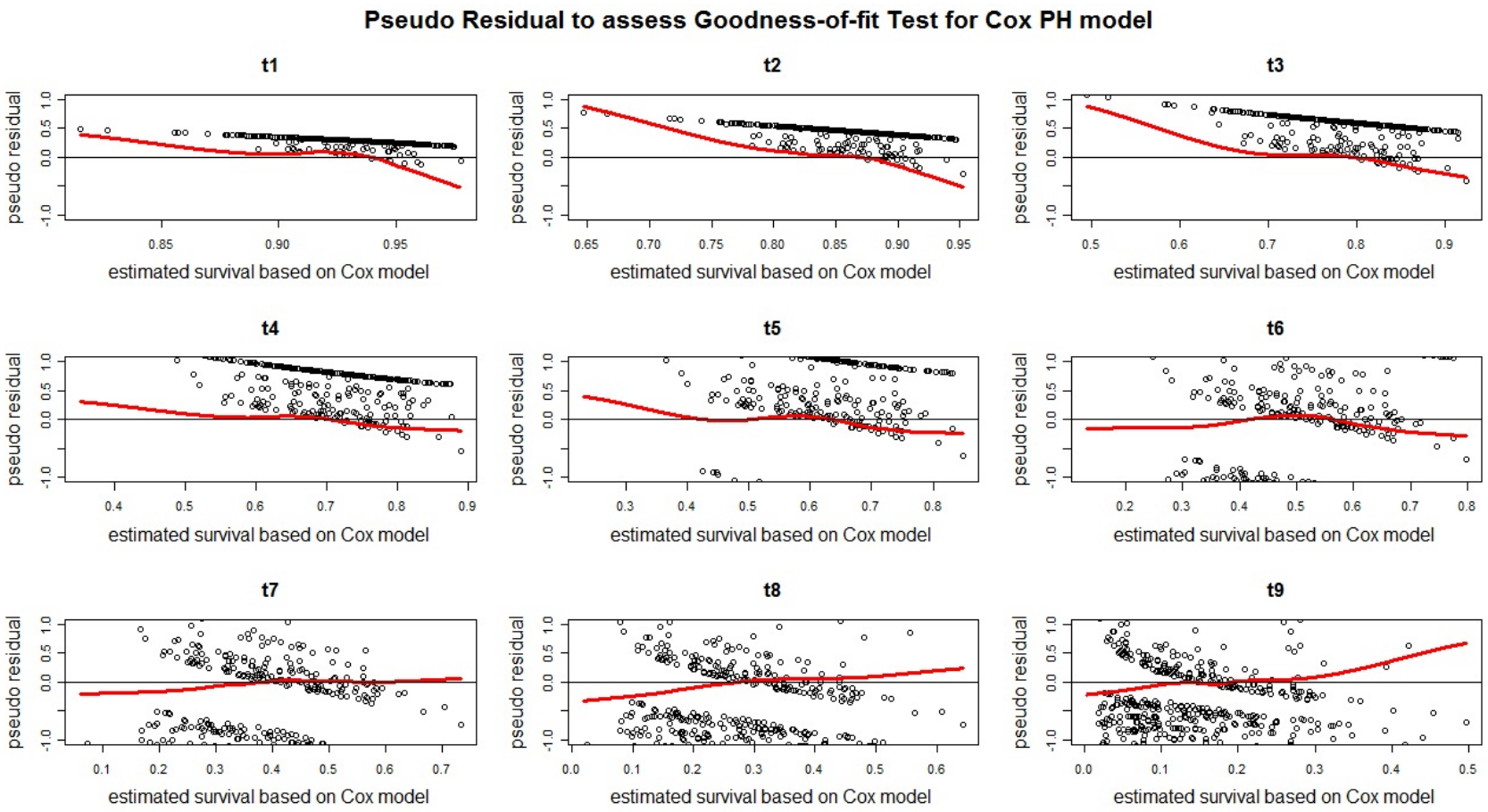

Figure 22: Simulation II-3 Pseudo residual vs. the estimated survival function based on Cox PH model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 34.6 % censoring

$$PseudoCox\ residual = \hat{S}_i(t) - \hat{S}_{Cox}(t|Z_i)$$

3.3.5 Simulation Study II-4 Heavy Censoring

Simulation II-4 was performed using survival data generated with $\theta = 2$, $\lambda = 1$, and heavy censoring (71%). The results are presented in Table 6 and Figure 23 trough Figure 28. The value of the test statistic for proportional hazards assumption is 5.427 (p = .023). This test statistic states that in true model, the effects of the covariate are varying over time.

The true regression coefficient, the estimated regression coefficient based on Gray's time-varying coefficients model, and the estimated regression coefficient based on Cox PH model at each 9 time knot are presented in Table 6. While the estimated regression coefficient based on Cox's model are constant over time as .310, the estimated regression coefficient based on Gray's model, (.150, .192, .212, .236, .293, .410, .540, .583, .565), and true regression coefficient, (.054, .130, .172, .216, .282, .304, .356, .416, .516), are changing over time. The estimated regression coefficients based on Gray's model is close to true regression coefficient except for early time points. The results of Table 6 were plotted in Figure 23 and Figure 24. The estimates of survival functions were calculated and plotted based on Gray's piecewise-costant time-varying model and Cox PH model at ($Z = 2, 4, 6, 8$) (Figure 25). Finally, pseudo residual plots and true-Gray residual plots along with smoothed average against the estimated survival rate based on fitted model at each 9 time point are presented in Figure 26, Figure 27, and Figure 28 to assess the goodness-of-fit for Gray's time-varying coefficients model and Cox PH model.

Table 6: Simulation II-4 Compare true value of covariate effect $\beta(t)$ to estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 66 % censoring

Time Knots	**.027**	**.065**	**.086**	**.108**	**.141**	**.152**	**.178**	**.208**	**.258**
True $\beta(t)$: θ x Time (θ=2)	**.054**	**.130**	**.172**	**.216**	**.282**	**.304**	**.356**	**.416**	**.516**
$\hat{\beta}(t)$ based on Gray's model	**.150**	**.192**	**.212**	**.236**	**.293**	**.410**	**.540**	**.583**	**.565**
$\hat{\beta}$ based on Cox model	**.326**	**.326**	**.326**	**.326**	**.326**	**.326**	**.326**	**.326**	**.326**

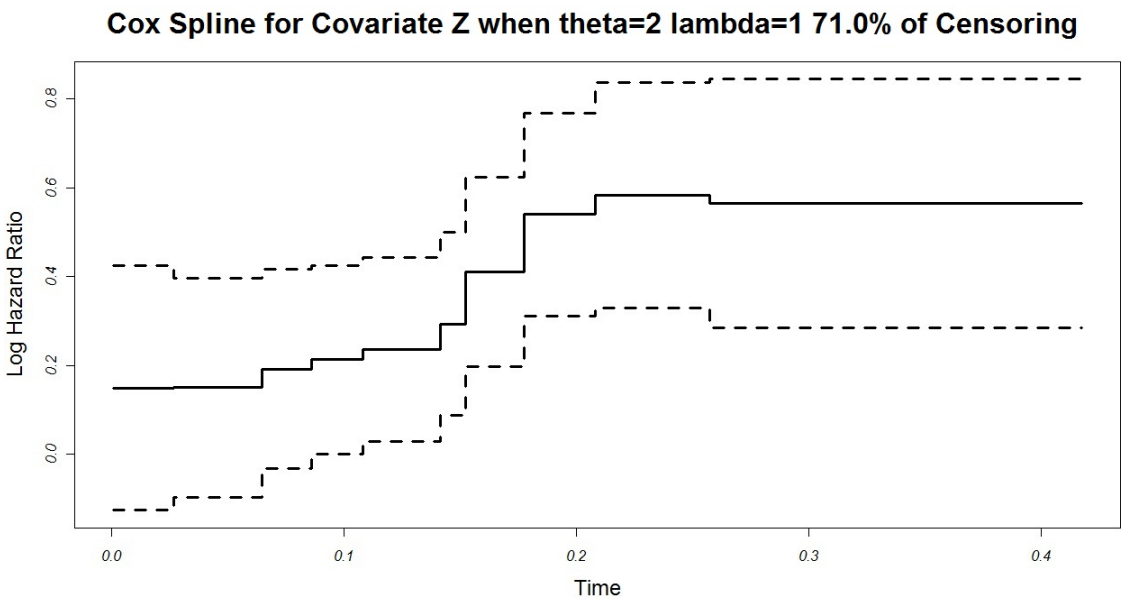

Figure 23: Simulation II-4 Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model (θ=2, λ=1, and 66% censoring)

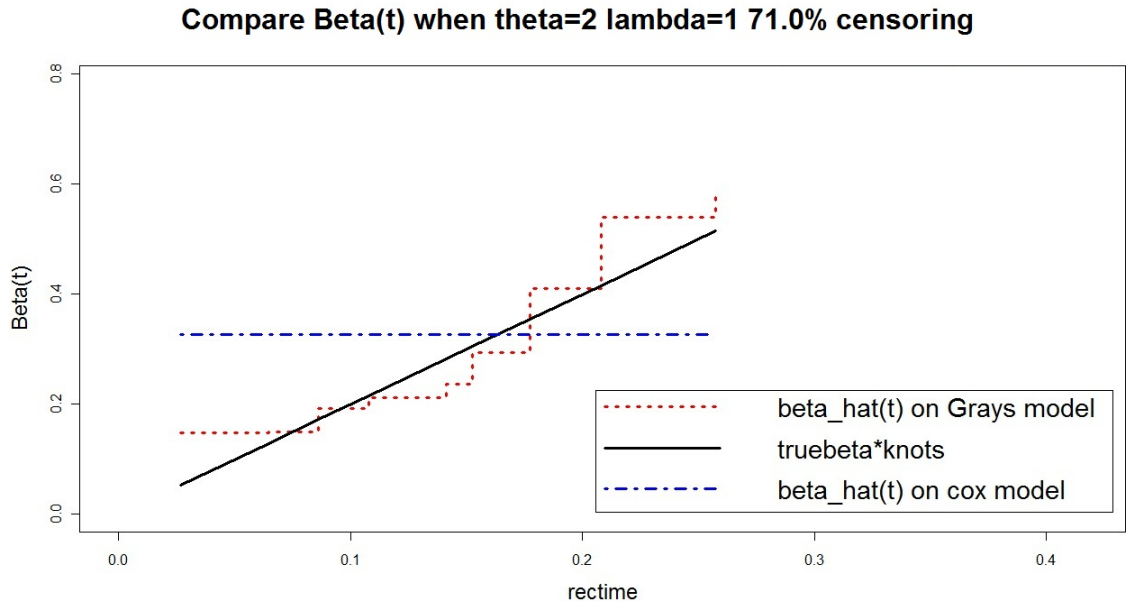

Figure 24: Simulation II-4 Plot of the true value of covariate effect $\beta(t)$, the estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model vs. time; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 66 % censoring

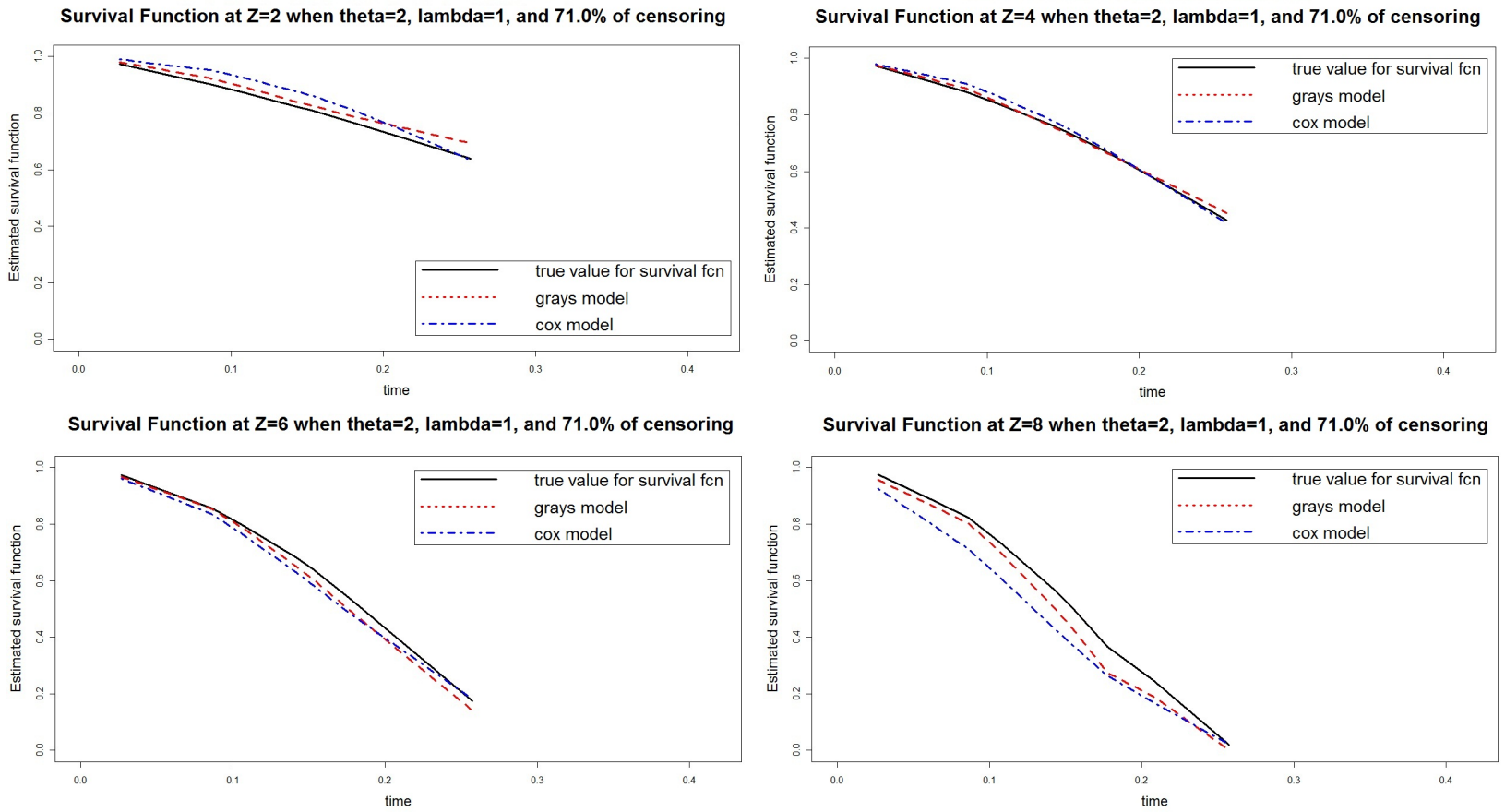

Figure 25: Simulation II-4 Plot of the estimated survival function based on Gray's time-varying coefficients and Cox PH model, and true survival function at (z=2,4,6, and 8): ; true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 66 % censoring

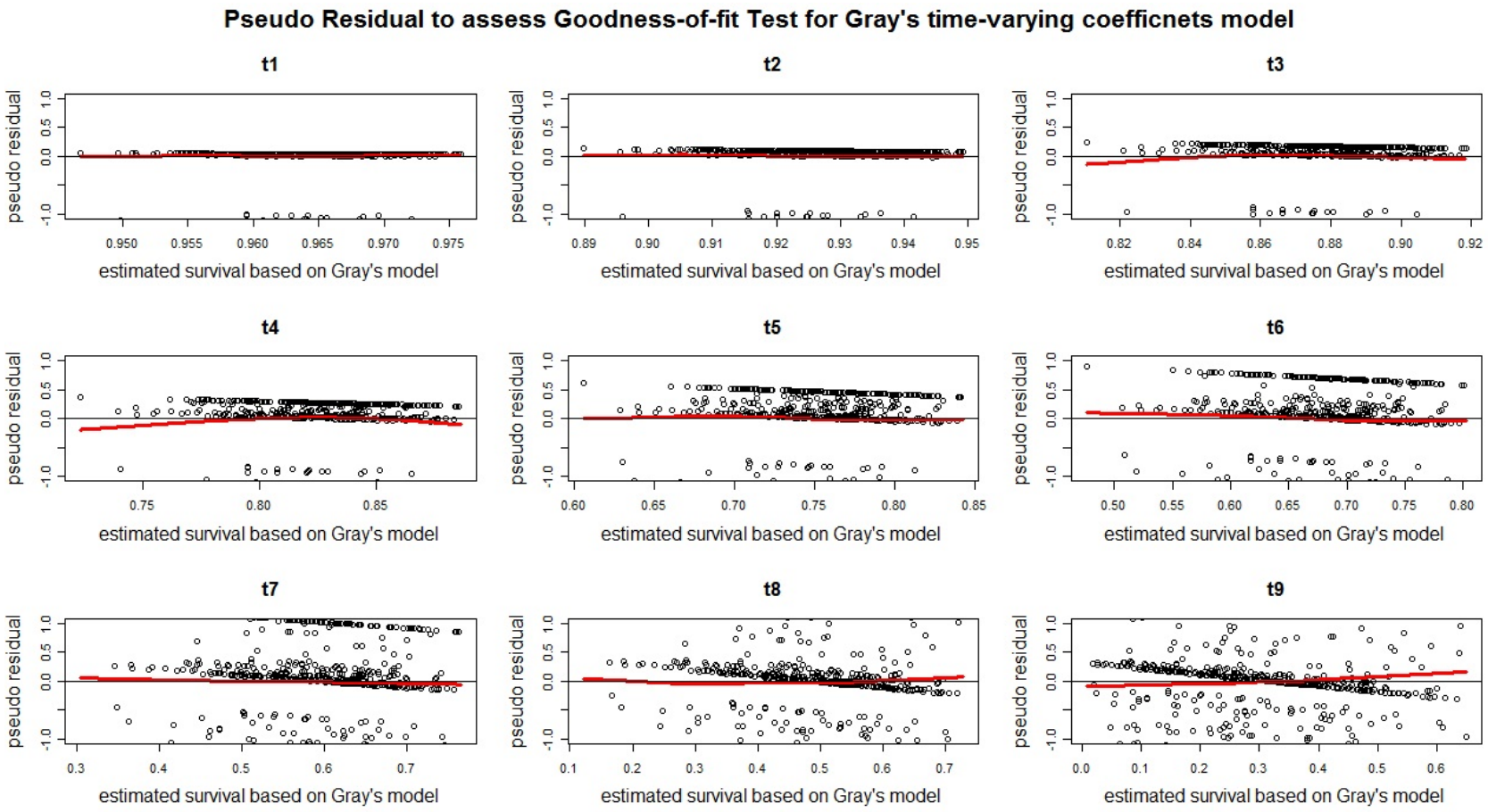

Figure 26: Simulation II-4 Pseudo residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 66 % censoring

$$PseudoGray\ residual = \hat{S}_i(t) - \hat{S}_{Gray}(t|Z_i)$$

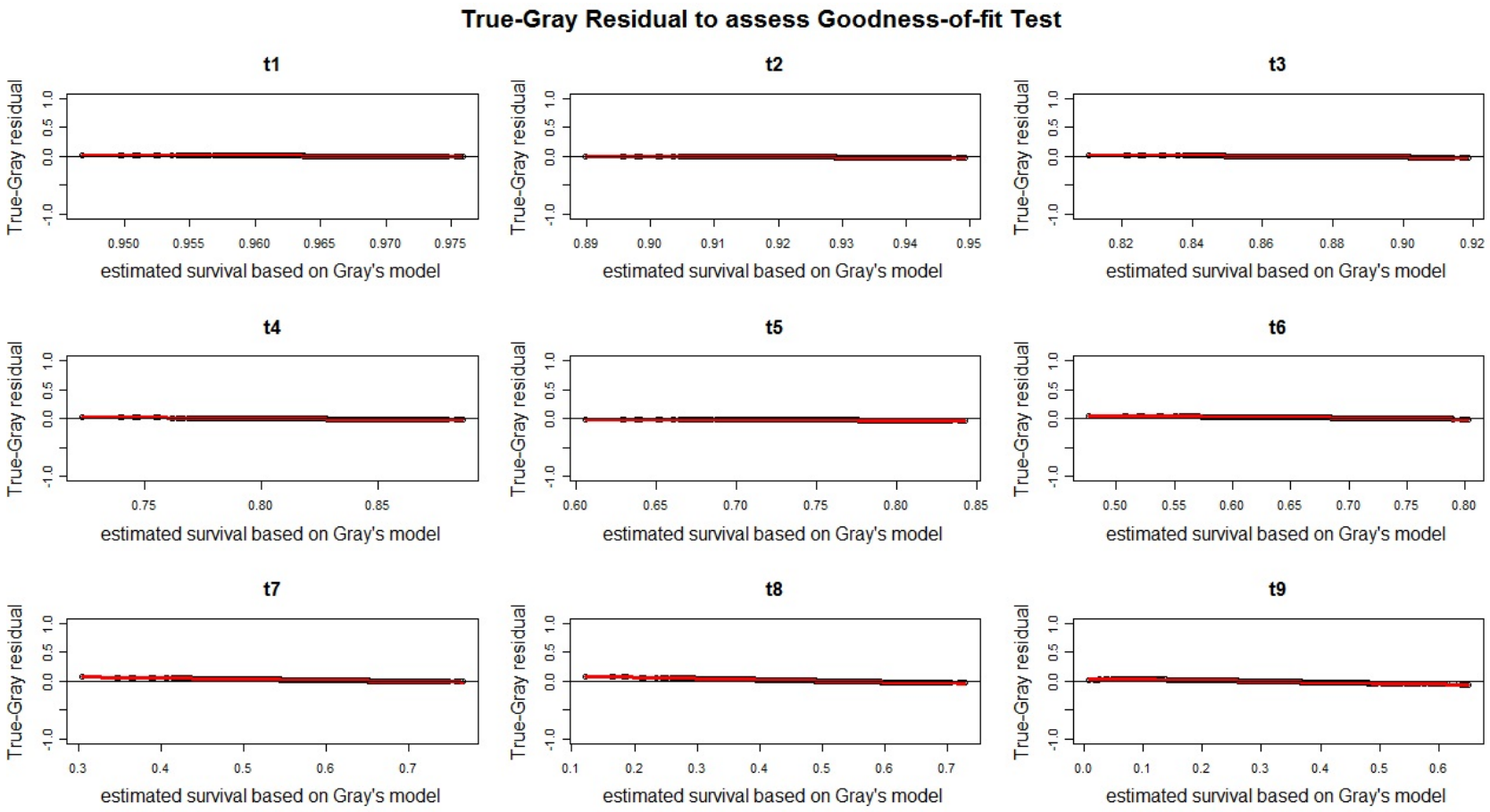

Figure 27: Simulation II-4 True-Gray residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 66 % censoring

$$TrueGray\ residual = S_{True}(t|Z_i) - \hat{S}_{Gray}(t|Z_i)$$

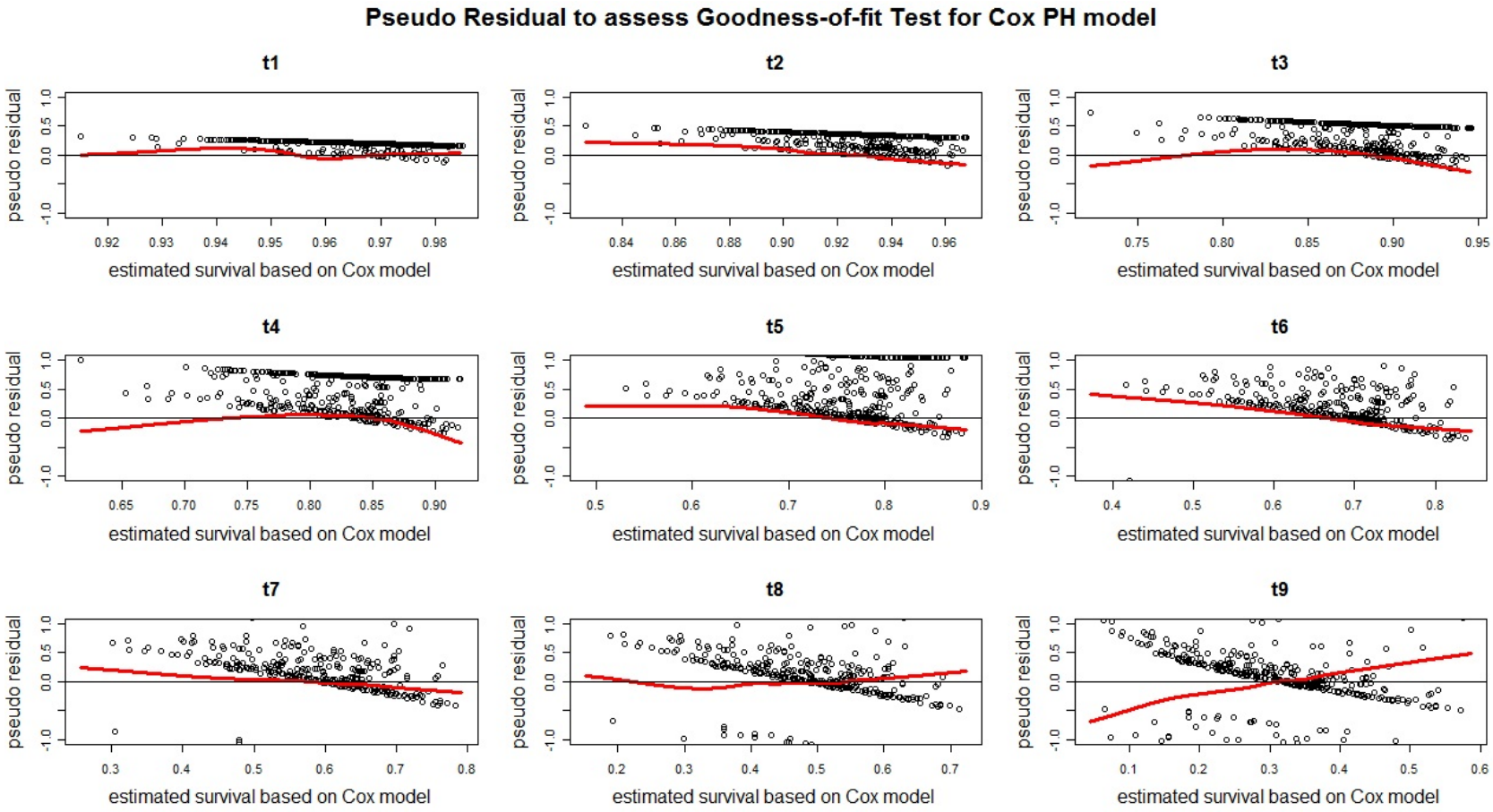

Figure 28: Simulation II-4 Pseudo residual vs. the estimated survival function based on Cox PH model at each time knot: true value of covariate effect $\beta(t) = 2t$, $\lambda = 1$, and 66 % censoring

$$PseudoCox\ residual = \hat{S}_i(t) - \hat{S}_{Cox}(t|Z_i)$$

3.4 RESULTS FOR DIFFERENT THETA, LAMBDA, AND PERCENTAGE OF CENSORING

Table 7 and Table 8 present the results of true value of survival function and estimated β based on Gray's piecewise-constant time-varying coefficient model and Cox proportional hazards ratio model for different theta, lambda, and censoring percentage.

3.5 CONCLUSION

In this simulation study, the sample size of 500 data, which contains $(T_i,\ \delta_i,\ Z_i)$ in, were generated via hazard function for time-varying model 2.11: $h(t) = h_0(t)\exp(\beta(t)Z)$ with $\beta(t) = 2t$, $\lambda = 1$, and varying of percentage of censoring. Nine time knots were predetermined to be evenly spaced using the 10 percentiles of event times.

The true covariate effect, the estimated covariate effect based on Gray's time-varying coefficients model and the estimated covariate effect based on Cox PH model were calculated and plotted against time. In the simulation I, the average of pseudo-observations, the estimated survival function based on Gray's time-varying coefficients model, and the estimated survival function based on Cox PH model were calculated for all individual at predetermined 9 time knots. The true survival function was also calculated by direct input of model setting. In the simulation II, the estimated survival function based on Gray's time-varying coefficients model and Cox PH model, and the true survival function were calculated for $(Z = 2,\ 4,\ 6,\ 8)$. Then, theses estimated survival functions were plotted against time. Finally, the residuals using pseudo-observation and true survival function for Gray's time-varying coefficients model and Cox PH model were calculated. To assess goodness-of-fit test using these residuals as a graphical tool, the residual plots along with smoothed average plots were conducted against the estimated survival rate based on fitted model at each 9 time knot.

Table 7: Compare true value of covariate effect $\beta(t)$ to estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model: $(\theta = 1,\ 2\ , 5)$,$(\lambda = .1,\ 1)$, and no censoring

θ	λ	Test Statistic (p-value) for Overall	Test Statistic (p-value) for Nonpro-portional	Time	1	2	3	4	5	6	7	8	9
1	1	53.49 (p<.001)	13.06 (p<.001)	True Beta: Theta x Time knots	.09	.18	.25	.30	.36	.43	.48	.55	.61
				Beta_hat (t) based on Gray's model	.06	.14	.25	.30	.39	.46	.51	.62	.72
				Beta_hat based on Cox model	.41	.41	.41	.41	.41	.41	.41	.41	.41
2	0.1	241.73 (p<.001)	23.61 (p<.001)	True Beta: Theta x Time knots	.57	.77	.94	1.04	1.16	1.26	1.39	1.54	1.70
				Beta_hat (t) based on Gray's model	.63	.72	.90	1.08	1.17	1.26	1.33	1.43	1.52
				Beta_hat based on Cox model	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08	1.08
	1	80.15 (p<.001)	17.66 (p<.001)	True Beta: Theta x Time knots	.16	.24	.33	.42	.49	.56	.62	.69	.81
				Beta_hat (t) based on Gray's model	.15	.22	.26	.32	.38	.50	.59	.63	.65
				Beta_hat based on Cox model	.44	.44	.44	.44	.44	.44	.44	.44	.44
5	0.1	231.42 (p<.001)	26.42 (p<.001)	True Beta: Theta x Time knots	.86	1.11	1.22	1.32	1.44	1.54	1.69	1.83	1.98
				Beta_hat (t) based on Gray's model	.88	1.04	1.16	1.26	1.37	1.50	1.51	1.62	1.65
				Beta_hat based on Cox model	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31
	1	129.36 (p<.001)	20.98 (p<.001)	True Beta: Theta x Time knots	.86	1.11	1.22	1.32	1.44	1.54	1.69	1.83	1.98
				Beta_hat (t) based on Gray's model	.88	1.04	1.16	1.26	1.37	1.50	1.51	1.62	1.65
				Beta_hat based on Cox model	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31	1.31

Table 8: Compare true value of covariate effect $\beta(t)$ to estimated covariate effect $\hat{\beta}(t)$ based on Gray's time-varying coefficient model and Cox PH model: $(\theta = 1,\ 2\ , 5)$,$(\lambda = .1,\ 1)$, and 20 % censoring

θ	λ	Test Statistic (p-value) for Overall	Test Statistic (p-value) for Nonpro-portional	Time	1	2	3	4	5	6	7	8	9
1	1	53.49 (p<.001)	13.06 (p<.001)	True Beta: Theta x Time knots	.10	.16	.21	.29	.35	.40	.44	.52	.63
				Beta_hat (t) based on Gray's model	.12	.18	.23	.29	.36	.47	.50	.56	.62
				Beta_hat based on Cox model	.38	.38	.38	.38	.38	.38	.38	.38	.38
2	0.1	241.73 (p<.001)	23.61 (p<.001)	True Beta: Theta x Time knots	.56	.73	.92	1.03	1.15	1.27	1.43	1.56	1.73
				Beta_hat (t) based on Gray's model	.72	.79	.93	.97	1.11	1.26	1.43	1.47	1.49
				Beta_hat based on Cox model	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
	1	80.15 (p<.001)	17.66 (p<.001)	True Beta: Theta x Time knots	.13	.27	.38	.44	.52	.58	.65	.75	.85
				Beta_hat (t) based on Gray's model	.32	.33	.36	.42	.46	.51	.61	.78	.92
				Beta_hat based on Cox model	.55	.55	.55	.55	.55	.55	.55	.55	.55
5	0.1	231.42 (p<.001)	26.42 (p<.001)	True Beta: Theta x Time knots	.81	1.03	1.67	1.29	1.39	1.53	1.66	1.82	2.03
				Beta_hat (t) based on Gray's model	.75	.84	1.04	1.16	1.27	1.35	1.44	1.59	1.78
				Beta_hat based on Cox model	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2
	1	129.36 (p<.001)	20.98 (p<.001)	True Beta: Theta x Time knots	.34	.47	.59	.70	.79	.88	.96	1.05	1.19
				Beta_hat (t) based on Gray's model	.37	.47	.57	.66	.84	.96	1.01	1.04	1.14
				Beta_hat based on Cox model	.77	.77	.77	.77	.77	.77	.77	.77	.77

In Table 1, Table 2, Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, the results represent that the estimated regression coefficient based on Gray's time-varying coefficient model is quite close to the true value regression coefficient. According to plot of log-hazard ratio and plot of estimated regression coefficient vs. time, (Figure 5, Figure 6, Figure 11, Figure 12, Figure 17, Figure 18, Figure 23, Figure 24), the effects of the covariate are changing over time, i.e., time-varying coefficients. The plots of the estimated survival function based on Gray's time-varying coefficients model and Cox PH model along with the true survival function are presented in Figure 7, Figure 13, Figure 19, Figure 25. Comparing these three plots, Gray's time-varying coefficients model is found to be a good fit in estimating survival function. The pseudo residual plots along with smoothed average against the estimated survival rate based on fitted model at each time point for Gray's model and Cox model are presented in Figure 8, Figure 10, Figure 14, Figure 16, Figure 20, Figure 22, Figure 26, Figure 28. Pseudo residuals for Gray's time-varying coefficients model are constant near 0 without any significant tendency or departure indicating a good shape of fit, while pseudo residuals for Cox PH model show tendencies and departures indicating lack of fit. The simulations were performed with varying percentage of censoring and the results from complete and right censored survival data are very similar. Also, the simulations were performed with different λ and different θ(Table 7 and Table 8). The results are very similar.

In conclusion, using Gray's time-varying coefficients model shows a good fit in estimating survival function when data violate the PH assumption. Also, the pseudo residual plot using pseudo-observations vs. estimated survival function based on Gray's time-varying coefficient model at each predetermined time knot is a useful graphical tool to assess goodness-of-fit test for Gray's time-varying coefficients model.

4.0 APPLICATION TO LIVER TRANSPLANTATION DATA

4.1 DATA DESCRIPTION

In this section, we apply our proposed method to assess the goodness-of-fit for a model that predicts the probability of posttransplant survival among children who were under the age of 12 years, had end-stage liver diseases, and underwent liver transplantation between January 2005 and June 2010. The data were derived from the United Network for Organ Sharing/Organ Procurement and Transplantation Network (UNOS/OPTN) containing 13 variables evaluated at the time of transplant and time from transplantation to posttransplant death date. If the recipients were alive at the study cutoff date, the time from transplantation to death were censored. After model selection, the final multivariable Gray's time-varying coefficients model includes the following covariates: age at transplant, multiorgan transplant or not (yes/no), disease diagnosis (acute liver diseases, sclerosis, autoimmune diseases, liver cancer, metabolic liver diseases, biliary atresia, and other liver diseases) and location right before transplantation (out of hospital, in hospital, and ICU).

4.2 METHODS AND RESULTS

Gray's model using piecewise-constant time-varying coefficients [33] assumes that $\beta(t)' = (\beta_1(t),\ \beta_2(t),\ \ldots,\ \beta_p(t))$, where $\beta_j(t)$ $(j = 1,\ \ldots,\ p)$ denotes the effects of the j^{th} risk factor, remains constant for the each time intervals between the selected time knots, $t \in [\tau_k,\ \tau_{k+1})$, and the regression coefficients $\beta_j(t)$ is allowed to change at the selected internal time knots, τ_k. In this analysis, nine equally distributed time knots were selected.

Then, proportional hazards assumption of each covariate was tested, and the estimated survival function was derived. The pseudo-observations were calculated for all subjects at nine selected time knots. The proportionality of each variable is tested using the Gray's test [34] provided in the R package cox.spline. According to the results, the covariate effects of age at time of transplant (p=.006), liver cancer diagnosis (p=.001), and at ICU before transplantation (p=.021) were significantly changing over time, while the covariate effects of other covariates were not statistically significantly changing over time (constant). The estimates of the regression coefficients at each selected time knots are listed in Table 9, Table 10, Table 11, and Table 12.

The plots of the log hazard ratio vs. time for age at transplantation, liver cancer diagnosis, and at ICU before transplantation are illustrated in Figure 29, Figure 30, and Figure 31. According to (2.14), the log hazard ratio is expressed by

$$\ln(HR) \;=\; \ln(\frac{h_0(t)\exp(\beta' Z)}{h_0(t)\exp(\beta' Z^*)}) = \ln(\exp(\beta(t)'(Z-Z^*))) = \beta(t)'(Z-Z^*)\,.$$

Therefore, the trend of $\hat{\beta}(t)$ can be checked from these plots. Figure 29, Figure 30, and Figure 31 show that the effects of these three covariates are changing over time, i.e., time-varying coefficients. The estimates of $\hat{\beta}(t)$ for all covariates are estimated based on Gray's time-varying coefficients model and Cox PH model, and plotted against time in Figure 32 through Figure 35. The estimated $\beta(t)$ for other covariates show constant over time, which are time-fixed coefficients.

The fitted model under Gray's time-varying coefficients model is stated in the final Gray's time-varying coefficients model 12.

Based on the final multivariable Gray's model, pseudo residuals were calculated and the plots of pseudo residuals against the estimated survival rates at each of the nine selected time point were constructed. These nine residual plots were then used to evaluate the goodness-of-fit of the final multivariable Gray's model (Figure 42 and Figure 43). For Gray's time-varying coefficients model, through the first time point and the sixth time point, pseudo residual plots are around zero, while pseudo residual plots show slight departure at small value of estimated survival rate through the seventh time point and the ninth time point.

Since the pseudo residual plots stays around zero at each time point without any significant departure or tendency, we can conclude that Gray's model shows a good fit in estimating survival function at each time point. The conclusion from this real liver transplantation data is that the final multivariable Gray's model shows a good fit in estimating posttransplant survival.

Table 9: The result of estimated time-varying covariate effect $\hat{\beta}(t)$ for age at transplant based on Gray's time-varying coefficients models and Cox PH models

variables	**age at transplant**								
time points	**56**	**84**	**140**	**176**	**282**	**360**	**516**	**767**	**1322**
estimated β based on Cox (p-vlue)	**0.007 (p=.055)**								
estimated β based on Gray's	**-0.0185**	**-0.0173**	**-0.011**	**-0.006**	**0.003**	**0.015**	**0.030**	**0.036**	**0.042**

Table 10: The result of estimated time-varying coefficient $\hat{\beta}(t)$ for liver cancer diagnosis based on Gray's time-varying coefficients models and Cox PH models

variables	**liver cancer diagnosis**								
time points	**56**	**84**	**140**	**176**	**282**	**360**	**516**	**767**	**1322**
estimated β based on Cox (p-vlue)	**1.246 (p<.001)**								
estimated β based on Gray's	**0.588**	**0.683**	**0.900**	**1.210**	**1.524**	**1.703**	**1.719**	**1.542**	**1.418**

Table 11: The result of estimated time-varying coefficient $\hat{\beta}(t)$ for at ICU before transplantation based on Gray's time-varying coefficients models and Cox PH models

variables	**at ICU before transplantation**								
time points	**56**	**84**	**140**	**176**	**282**	**360**	**516**	**767**	**1322**
estimated β based on Cox (p-vlue)	**.0673 (p<.001)**								
estimated β based on Gray's	**0.941**	**0.906**	**0.831**	**0.682**	**0.586**	**0.504**	**0.390**	**0.322**	**0.421**

Table 12: The result of estimated time-constant coefficients $\hat{\beta}(t)$'s for multiorgan transplant or not (yes/no), disease diagnosis (acute liver diseases, sclerosis, autoimmune diseases, metabolic liver diseases, and biliary atresia) and location right before transplantation (out of hospital) based on Gray's time-varying coefficients models and Cox PH models

		Diagnosis group					**Lpcation**
variables	**multiorgan transplant or not (yes/No)**	**acute liver diseases diagnosis**	**sclerosis diagnosis**	**autoimmune diseases diagnosis**	**metabolic liver diseases diagnosis**	**biliary atresia diagnosis**	**out of hospital before transplan-tation**
estimated β based on Cox (p-vlue)	**1.920 (p<.001)**	**.092 (p=.69)**	**.027 (p=.96)**	**.044 (p=.89)**	**.019 (p=.93)**	**.036 (p=.84)**	**.137 (p=.36)**
estimated β based on Gray's	**1.914**	**0.109**	**0.03**	**0.045**	**0.04**	**0.03**	**0.146**

Final Gray's time-varying coefficients model:

$h(t|Z) = h_o(t)\exp(\beta_1(t) * age\,at\,transplantation + \beta_2(t) * liver\,cancer\,diagnosis + \beta_3(t) * at\,ICU\,before\,transplantation + \beta_4 * multiorgan\,transplant\,or\,not + \beta_5 * acute\,liver\,diseases + \beta_6 * sclerosis + \beta_7 * autoimmune\,diseases + \beta_8 * metabolic\,liver\,diseases + \beta_9 * biliary\,atresia + \beta_{10} * out\,of\,hospital$

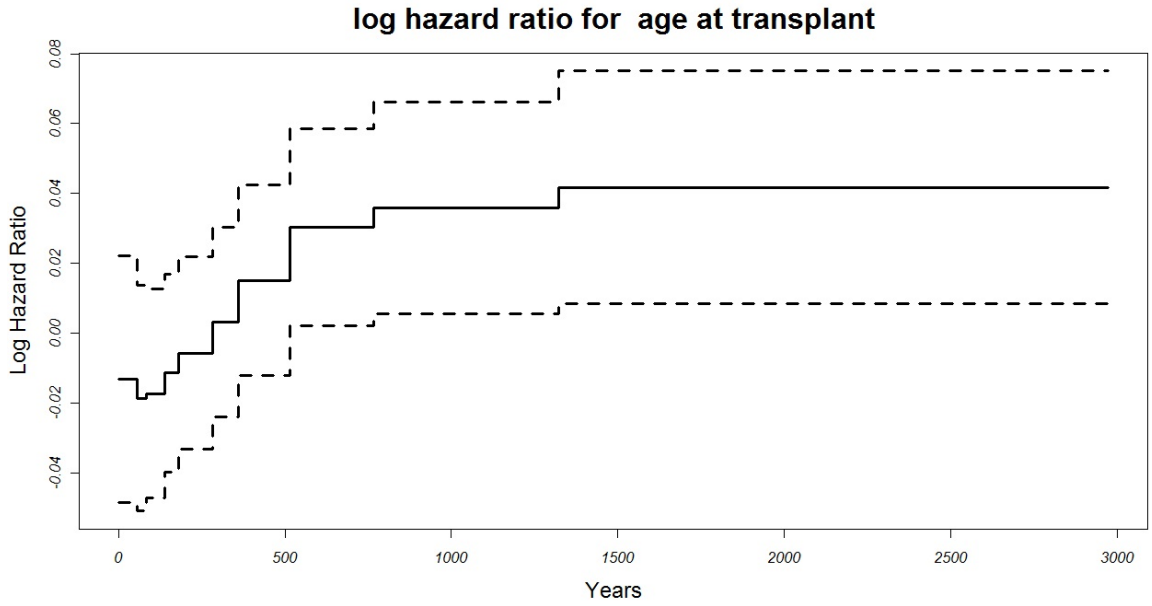

Figure 29: Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model for age at transplant

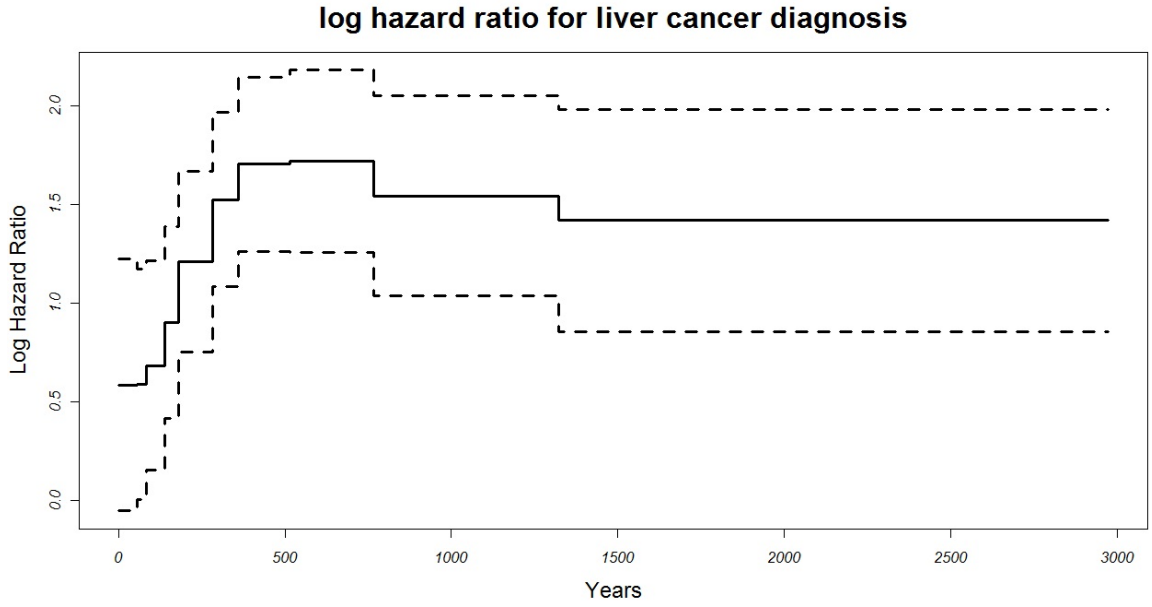

Figure 30: Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model for liver cancer diagnosis

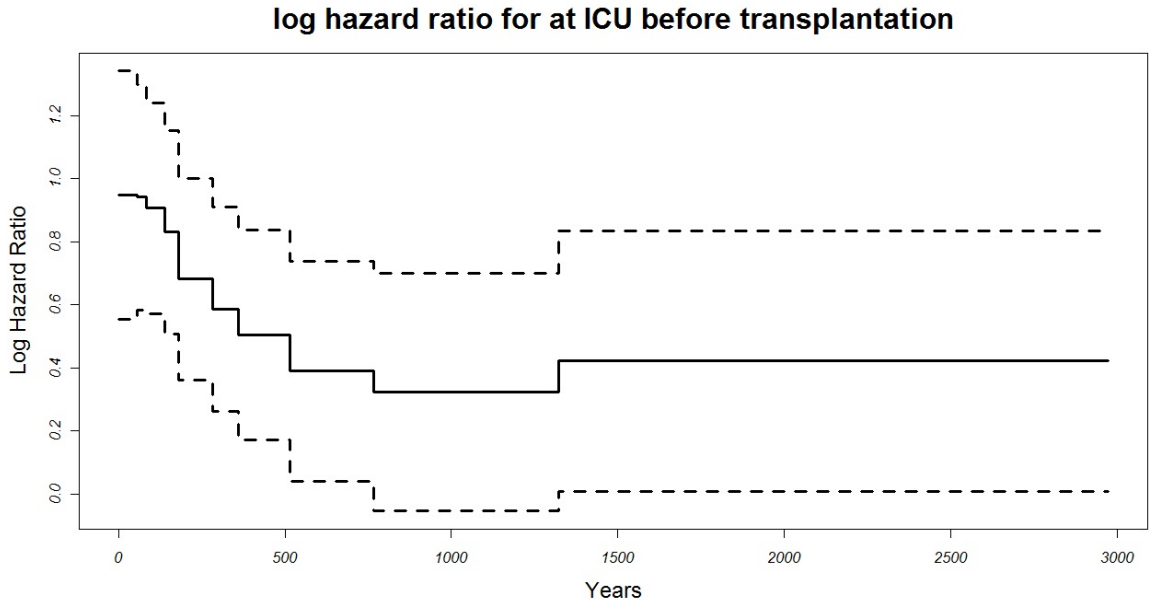

Figure 31: Plot of log hazard ratio vs. time for the covariate Z based on Gray's time-varying coefficients models and Cox PH model for at ICU before transplantation

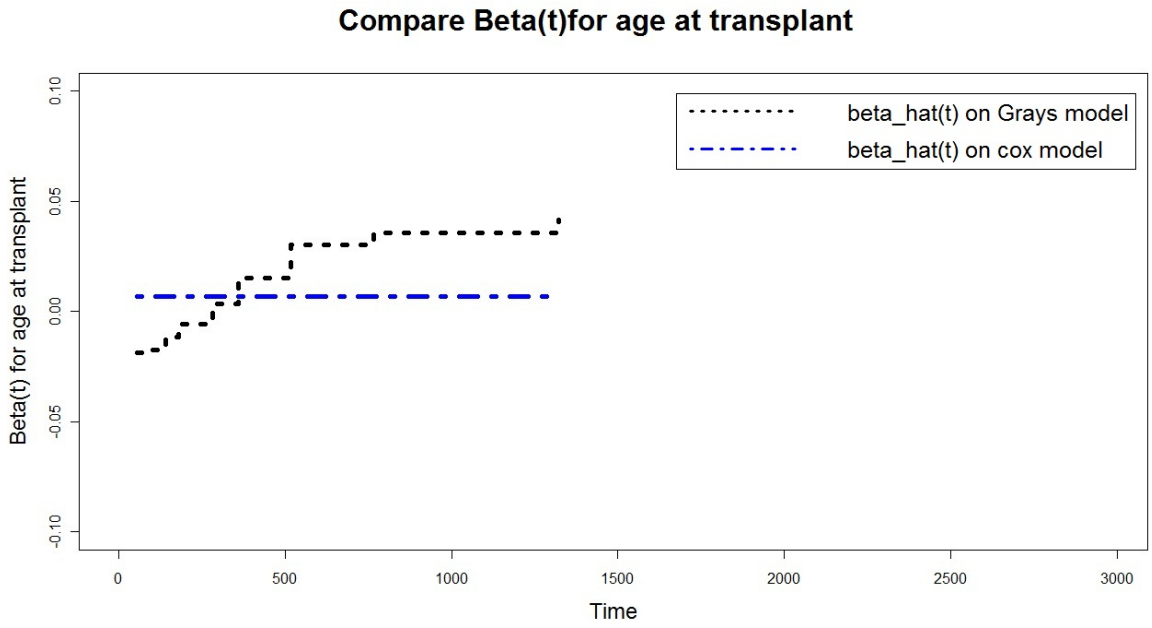

Figure 32: Plot of the estimated effect of covaraite age at transplant, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

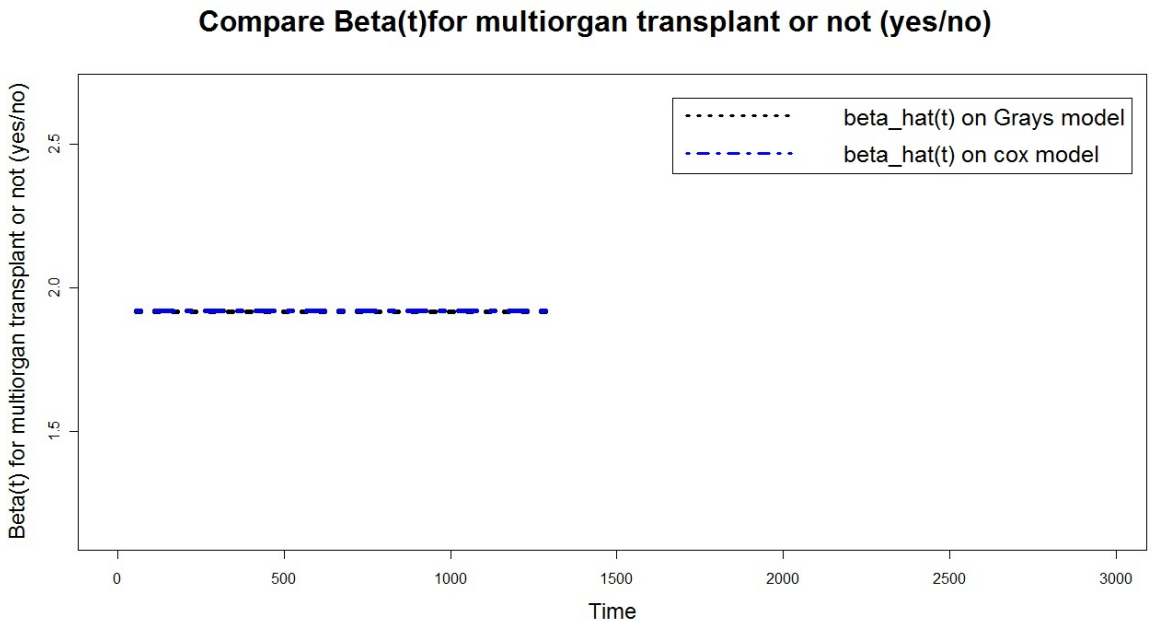

Figure 33: Plot of the estimated effect of covaraite multiorgan transplant or not (yes/no), $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. timel

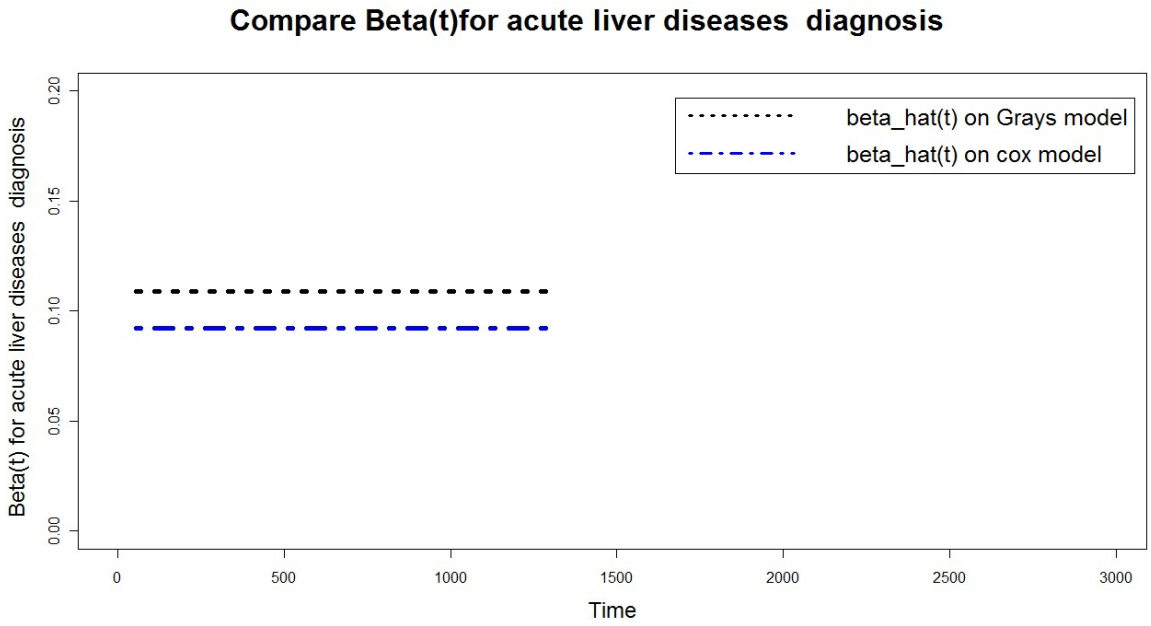

Figure 34: Plot of the estimated effect of covaraite acute liver diseases diagnosis, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

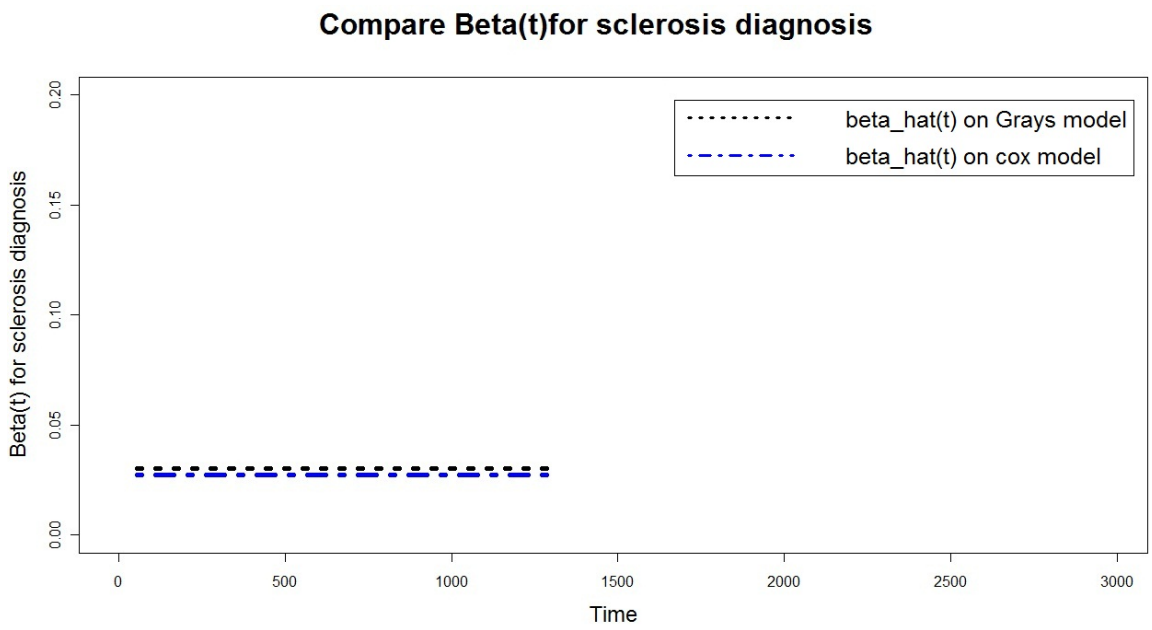

Figure 35: Plot of the estimated effect of covaraitesclerosis diagnosis, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

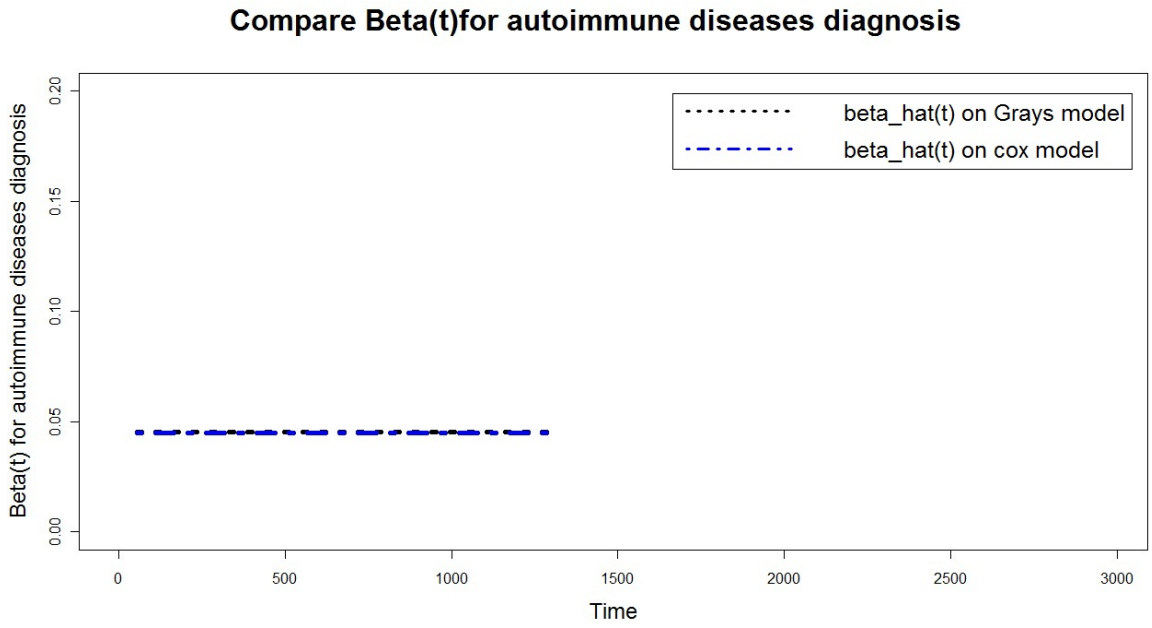

Figure 36: Plot of the estimated effect of covaraite autoimmune diseases diagnosis, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

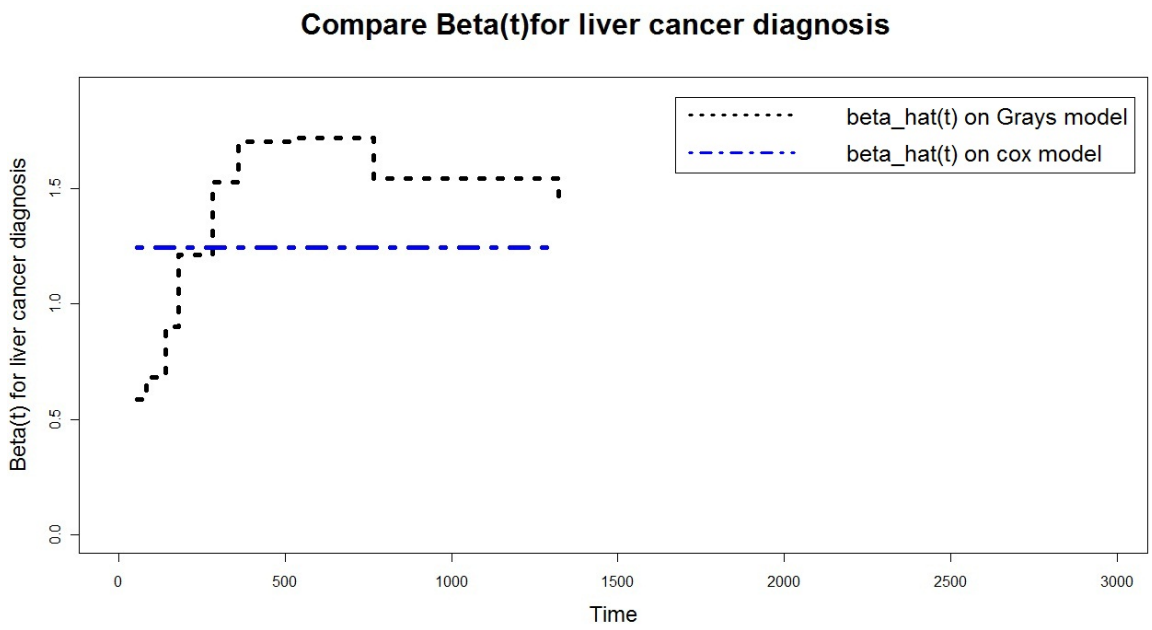

Figure 37: Plot of the estimated effect of covaraite liver cancer diagnosis, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

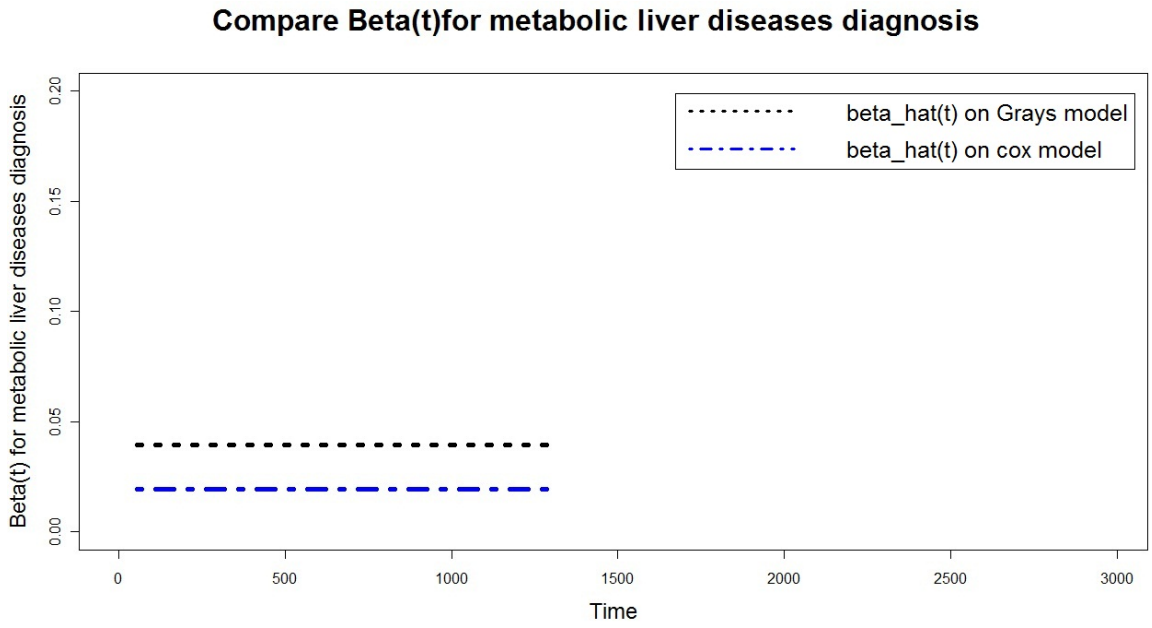

Figure 38: Plot of the estimated effect of covaraite metabolic liver diseases diagnosis, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

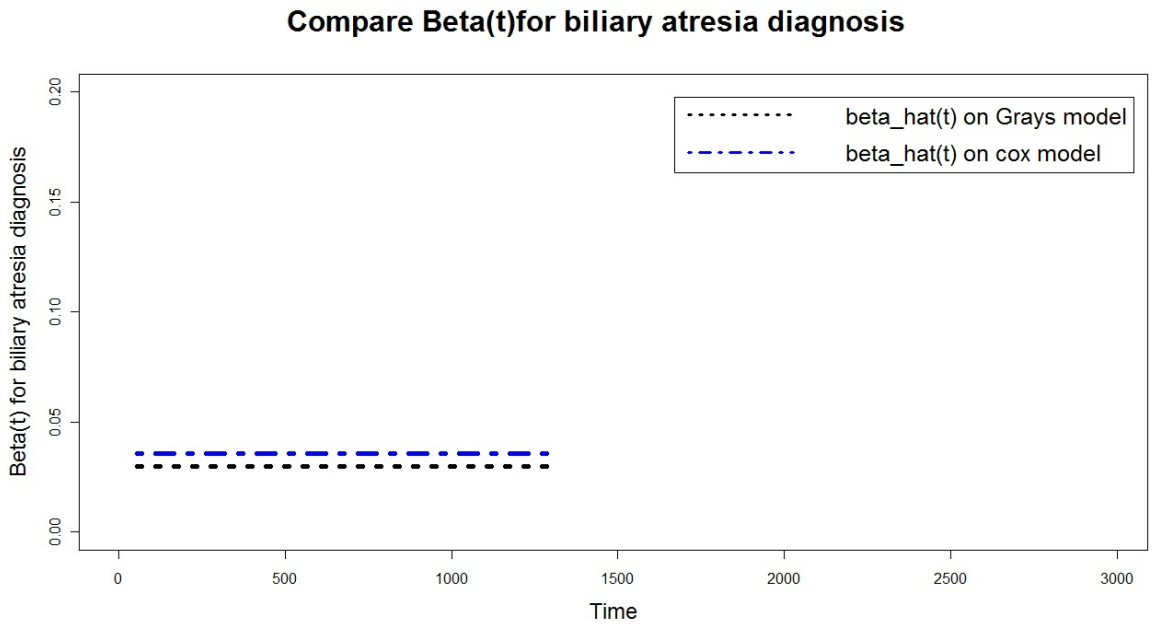

Figure 39: Plot of the estimated effect of covaraite biliary atresia diagnosis, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

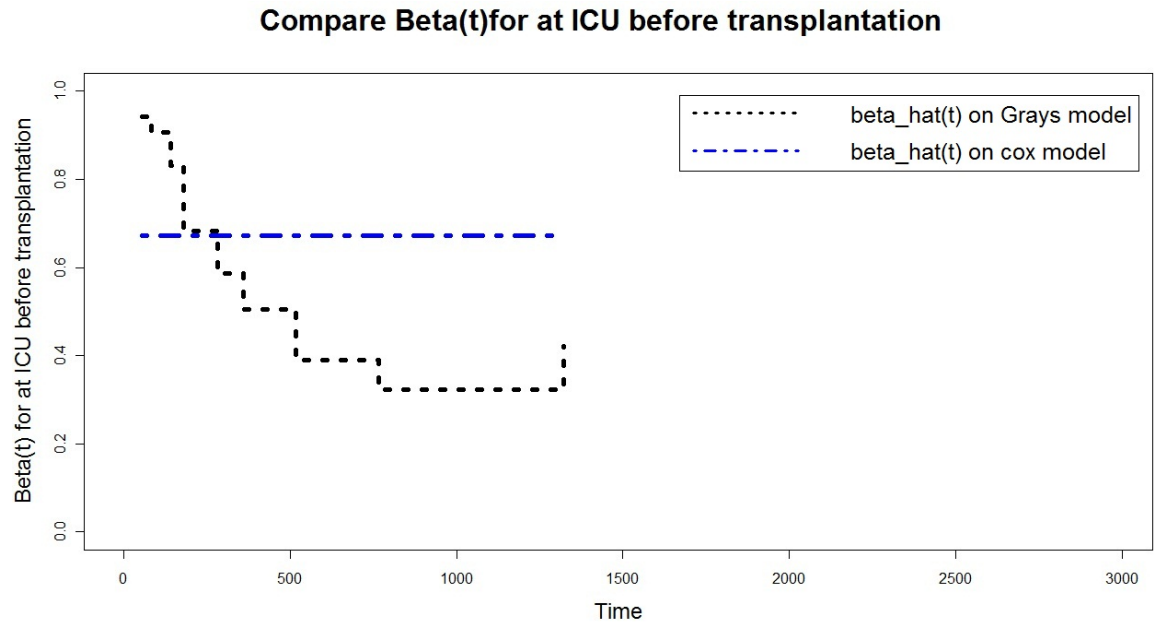

Figure 40: Plot of the estimated effect of covaraite at ICU before transplantation, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs. time

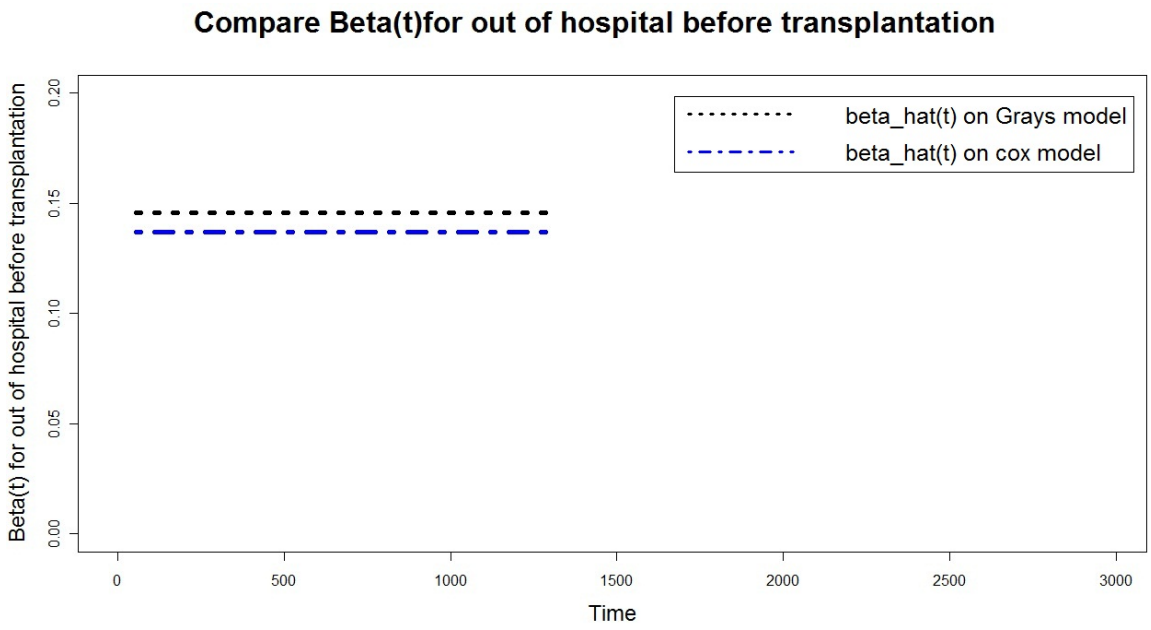

Figure 41: Plot of the estimated effect of covaraitef out of hospital before transplantation, $\hat{\beta}(t)$, based on Gray's time-varying coefficient model and Cox PH model vs.time

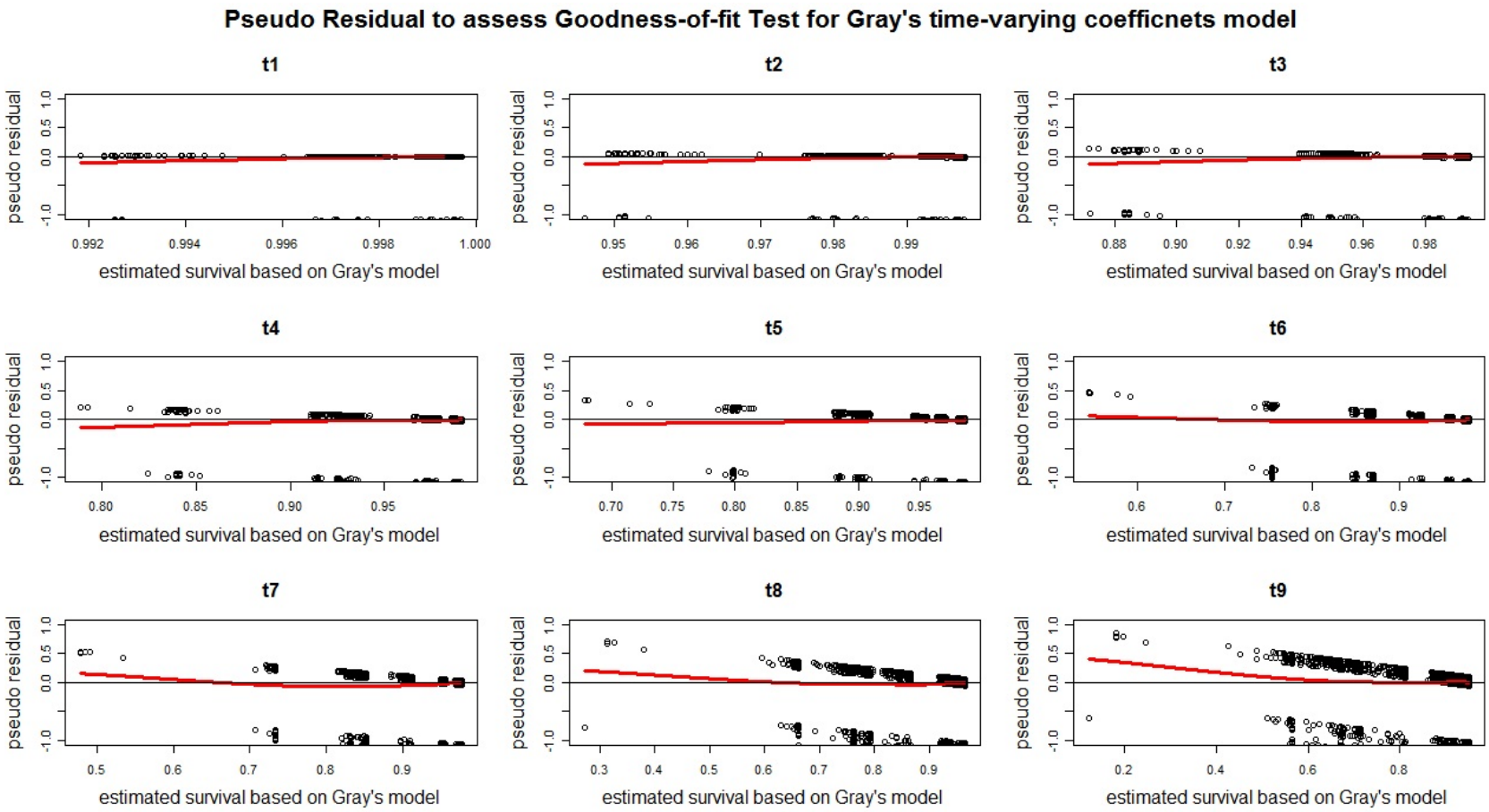

Figure 42: Pseudo residual vs. the estimated survival function based on Gray's time-varying coefficients model at each time knot

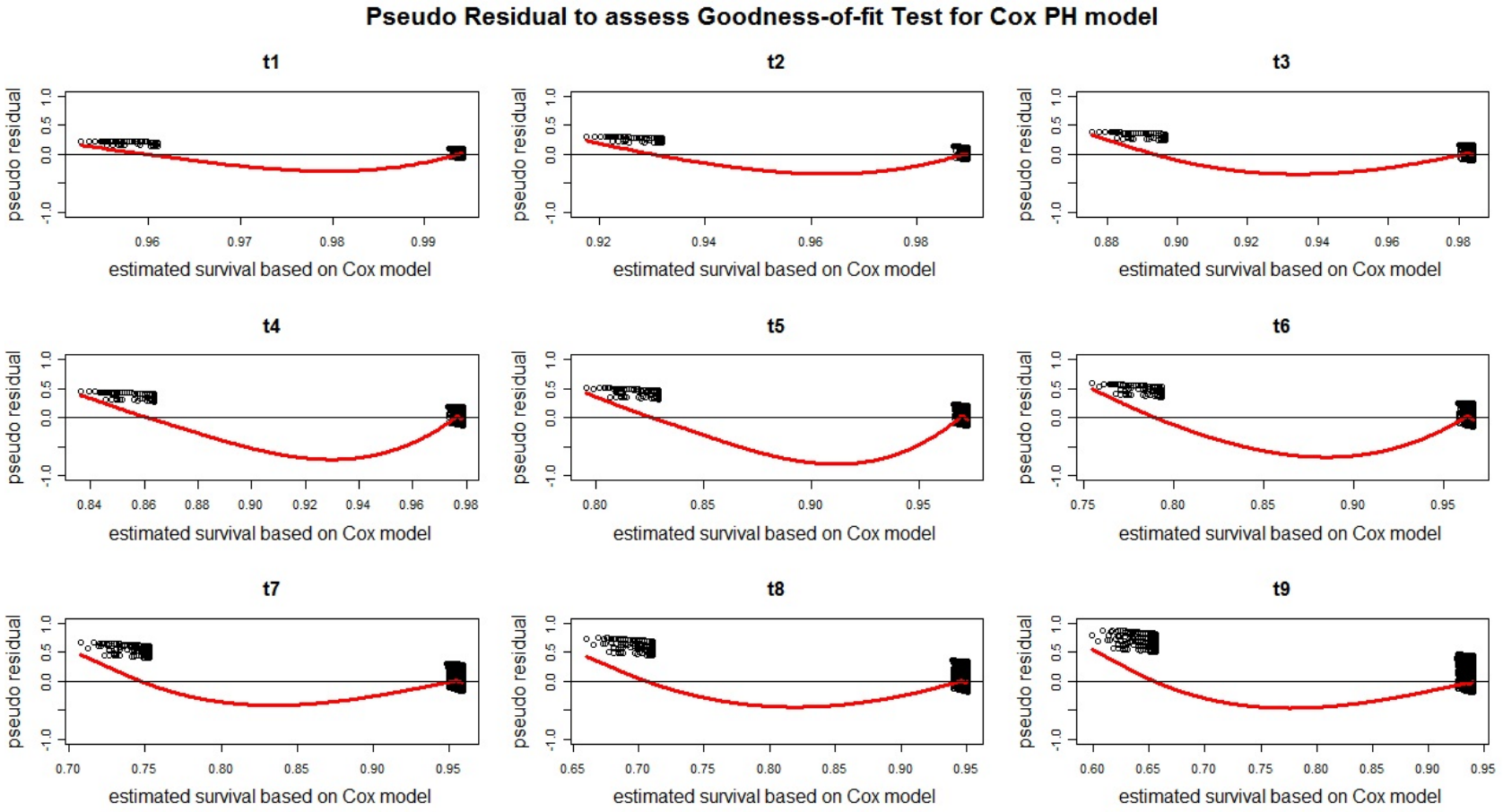

Figure 43: Pseudo residual vs. the estimated survival function based on Cox PH model at each time knot

5.0 CONCLUSION AND DISCUSSION

The Cox proportional hazards model[18] has become the standard for analyzing the effects of covariates for time-to-event data. However, in practical situations the assumption of proportional hazards often are not met, and time-varying effect of a covariate, $\beta(t)$, is meaningful and clinically important. Among several methods proposed to estimate time-varying effect of a covariate via a time varying coefficient $\beta(t)$, Gray's piecewise-constant time-varying coefficients model [33] was focused in this study. Currently, there is no method available to assess the overall goodness-of-fit for Gray's model. In this study, we proposed a goodness-of-fit method for Gray's model based on pseudo-observations [7], [8].

One way for assessing the model goodness-of-fit is via residual plots. In survival analysis, this method is complicated when data involve censoring. There are several reasons why the graphical methods are difficult for right censored data. One of the reasons is because the plots require the incorporation of one more dimension, which is the censoring information. Because censoring time is incompletely observed, it can result in an incorrect relationship between the survival times and the covariates. In this case, pseudo-observations can be used to replace the incomplete data due to censoring. In this study, pseudo-observations were calculated for all individuals at each predetermined time point. Pseudo residual plots based on pseudo-observations were then used as a graphical diagnostic tool to examine a model fit. In practice, because many points are overlapped on the residual plots, it is difficult to evaluate the trends from the plots. To resolve this issue, we suggested to plot smoothed averages along with the residual plots. Under proper model, there is no trend between pseudo residuals and estimated survival functions at any time point. If residual plots show some tendency or departure, it means that there exists lack of fit. Our simulation under various percentages of censoring gave very similar results.

Several notes are worth of attention. Censoring must be independent with the event time and with the covariates in the model to estimate unbiased pseudo-observation. The choice of the number of time points is not a critical issue since there is almost no difference between the results based on different numbers of time points. However, some departures may be presented in early and late time points. These are because either too small or too large numbers of events happened in a certain time interval to the individuals within certain ranges of the covariates [73]. Therefore, the plots in early or late time points should be interpreted with caution. The residual plots give a good initial tool to evaluate the model fit, but numerical tests should be supported for more accurate test result.

BIBLIOGRAPHY

[1] P.K. Andersen. Testing goodness of fit of cox's regression and life model. *Biometrics*, 38(1):67–77, 1982.

[2] PK Andersen, O. Borgan, RD Gill, and N. Keiding. Statistical models based on counting processes, 1992.

[3] P.K. Andersen, C.T. Ekstrøm, J.P. Klein, Y. Shu, and M.J. Zhang. A class of goodness of fit tests for a copula based on bivariate right-censored data. *Biometrical Journal*, 47(6):815–824, 2005.

[4] P.K. Andersen and R.D. Gill. Cox's regression model for counting processes: a large sample study. *The Annals of Statistics*, 10(4):1100–1120, 1982.

[5] P.K. Andersen, M.G. Hansen, and J.P. Klein. Regression analysis of restricted mean survival time based on pseudo-observations. *Lifetime Data Analysis*, 10(4):335–350, 2004.

[6] P.K. Andersen and J.P. Klein. Regression analysis for multistate models based on a pseudo-value approach, with applications to bone marrow transplantation studies. *Scandinavian Journal of Statistics Theory and Applications*, 34(1):3–16, 2007.

[7] P.K. Andersen, J.P. Klein, and S. Rosthøj. Generalised linear models for correlated pseudo-observations, with applications to multi-state models. *Biometrika*, 90(1):15–27, 2003.

[8] P.K. Andersen and M. Pohar Perme. Pseudo-observations in survival analysis. *Statistical Methods in Medical Research*, 19(1):71–99, 2010.

[9] E. Arjas. A graphical method for assessing goodness of fit in cox's proportional hazards model. *Journal of the American Statistical Association*, 83(401):204–212, 1988.

[10] R. Bender, T. Augustin, and M. Blettner. Generating survival times to simulate cox proportional hazards models. *Statistics in Medicine*, 24(11):1713–1723, 2005.

[11] E.H. Blackstone, D.C. Naftel, and M.E. Turner. The decomposition of time-varying hazard into phases, each incorporating a separate stream of concomitant information. *Journal of the American Statistical Association*, 81(395):615–624, 1986.

[12] W. Boehm. Inserting new knots into b-spline curves. *Computer-Aided Design*, 12(4):199–201, 1980.

[13] P.L. Bonate. Pharmacokinetic-pharmacodynamic modeling and simulation. *American Journal of Pharmaceutical Education*, 71(1):18, 2007.

[14] IS Borisov. Approximation of empirical fields, constructed with respect to vector observations with dependent components. *Siberian Mathematical Journal*, 23(5):615–623, 1982.

[15] N. Breslow and J. Crowley. A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, 2(3):437–453, 1974.

[16] Y.H.J. Chen and G.H.F. Liu. A note on the estimate of treatment effect from a cox regression model when the proportionality assumption is violated. *Communications in Statistics-Theory and Methods*, 35(3):521–526, 2006.

[17] A. Ciampi and J. Etezadi-Amoli. A general model for testing the proportional hazards and the accelerated failure time hypotheses in the analysis of censored survival data with covariates. *Communications in Statistics-Theory and Methods*, 14(3):651–667, 1985.

[18] D.R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(2):187–220, 1972.

[19] D.R. Cox and E.J. Snell. A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(2):248–275, 1968.

[20] J. Crowley and BE Storer. Comment on "a reanalysis of the stanford heart transplant data" by m. aitkin, n. laird, and b. francis. *Journal of the American Statistical Association*, 78:277–281, 1983.

[21] C. De Boor. A practical guide to splines, volume 27 of applied mathematical sciences, 1978.

[22] R.L. Eubank and C.H. Spiegelman. Testing the goodness of fit of a linear model via nonparametric regression techniques. *Journal of the American Statistical Association*, 85(410):387–392, 1990.

[23] R.Z. Ezzeddine. A comparison of test statistics for proportionality of the hazards in the cox regression model. 2003.

[24] L.D. Fisher and DY Lin. Time-dependent covariates in the cox proportional-hazards regression model. *Annual Review of Public Health*, 20(1):145–157, 1999.

[25] M.H. Gail and J.L. Gastwirth. A scale-free goodness-of-fit test for the exponential distribution based on the lorenz curve. *Journal of the American Statistical Association*, 73(364):787–793, 1978.

[26] D. Gamerman and M. West. An application of dynamic survival models in unemployment studies. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 36(2/3):269–274, 1987.

[27] R. Gill and M. Schumacher. A simple test of the proportional hazards assumption. *Biometrika*, 74(2):289–300, 1987.

[28] R.D. Gill. Understanding cox's regression model: a martingale approach. *Journal of the American Statistical Association*, 79(386):441–447, 1984.

[29] R. Giorgi, M. Abrahamowicz, C. Quantin, P. Bolard, J. Esteve, J. Gouvernet, and J. Faivre. A relative survival regression model using b-spline functions to model non-proportional hazards. *Statistics in Medicine*, 22(17):2767–2784, 2003.

[30] P.M. Grambsch and T.M. Therneau. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3):515–526, 1994.

[31] F. Graw, T.A. Gerds, and M. Schumacher. On pseudo-values for regression analysis in competing risks models. *Lifetime Data Analysis*, 15(2):241–255, 2009.

[32] R.J. Gray. Some diagnostic methods for cox regression models through hazard smoothing. *Biometrics*, 46(1):93–102, 1990.

[33] R.J. Gray. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*, 87(420):942–951, 1992.

[34] R.J. Gray. Spline-based tests in survival analysis. *Biometrics*, 50(3):640–652, 1994.

[35] R.J. Gray and D.A. Pierce. Goodness-of-fit tests for censored survival data. *The Annals of Statistics*, 13(2):552–563, 1985.

[36] U. Halekoh, S. Højsgaard, and J. Yan. The r package geepack for generalized estimating equations. *Journal of Statistical Software*, 15(2):1–11, 2006.

[37] T. Hastie and R. Tibshirani. Varying-coefficient models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 55(4):757–796, 1993.

[38] K.R. Hess. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in Medicine*, 13(10):1045–1062, 1994.

[39] J.L. Horowitz and G.R. Neumann. A generalized moments specification test of the proportional hazards model. *Journal of the American Statistical Association*, 87(417):234–240, 1992.

[40] D.W. Hosmer, T. Hosmer, S. Le Cessie, and S. Lemeshow. A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine*, 16(9):965–980, 1997.

[41] J.D. Kalbfleisch, R.L. Prentice, and JD Kalbfleisch. *The statistical analysis of failure time data*, volume 5. Wiley New York:, 1980.

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

[43] J. Kasal, Z. Jovanovic, G. Clermont, L.A. Weissfeld, V. Kaplan, R.S. Watson, and D.C. Angus. Comparison of cox and gray's survival models in severe sepsis*. *Critical Care Medicine*, 32(3):700–707, 2004.

[44] R. Kay. Proportional hazard regression models and the analysis of censored survival data. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 26(3):227–237, 1977.

[45] J.P. Klein, M. Gerster, P.K. Andersen, S. Tarima, and M.P. Perme. Sas and r functions to compute pseudo-values for censored data regression. *Computer methods and programs in biomedicine*, 89(3):289–300, 2008.

[46] J.P. Klein and M.L. Moeschberger. *Survival analysis: techniques for censored and truncated data*. Springer Verlag, 2003.

[47] D.G. Kleinbaum and M. Klein. *Survival analysis: a self-learning text*. Springer Verlag, 2005.

[48] S.W. Lagakos. The graphical evaluation of explanatory variables in proportional hazard regression models. *Biometrika*, 68(1):93–98, 1981.

[49] M. LeBlanc and J. Crowley. Adaptive regression splines in the cox model. *Biometrics*, 55(1):204–213, 1999.

[50] G. Li and Y. Sun. A simulation-based goodness-of-fit test for survival data. *Statistics & Probability Letters*, 47(4):403–410, 2000.

[51] Y.H. Li, J.P. Klein, and M.L. Moeschberger. Effects of model misspecification in estimating covariate effects in survival analysis for small sample sizes. *Computational Statistics & Data analysis*, 22(2):177–192, 1996.

[52] K.Y. Liang and S.L. Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986.

[53] D.Y. Lin. Goodness-of-fit analysis for the cox regression model based on a class of parameter estimators. *Journal of the American Statistical Association*, 86(415):725–728, 1991.

[54] DY Lin and LJ Wei. Goodness-of-fit tests for the general cox regression model. *Statistica Sinica*, 1(1):1–17, 1991.

[55] D.Y. Lin, L.J. Wei, and Z. Ying. Checking the cox model with cumulative sums of martingale-based residuals. *Biometrika*, 80(3):557–572, 1993.

[56] D.Y. Lin and Z. Ying. Semiparametric analysis of general additive-multiplicative hazard models for counting processes. *The Annals of Statistics*, 23(5):1712–1734, 1995.

[57] T. Martinussen, T.H. Scheike, and I.M. Skovgaard. Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. *Scandinavian Journal of Statistics*, 29(1):57–74, 2002.

[58] S. May and D.W. Hosmer. A simplified method of calculating an overall goodness-of-fit test for the cox proportional hazards model. *Lifetime Data Analysis*, 4(2):109–120, 1998.

[59] I.W. McKeague and K.J. Utikal. Goodness-of-fit tests for additive hazards and proportional hazards models. *Scandinavian Journal of Statistics*, 18(3):177–195, 1991.

[60] R. Miller and J. Halpern. Regression with censored data. *Biometrika*, 69(3):521–531, 1982.

[61] R.G. Miller. The jackknife-a review. *Biometrika*, 61(1):1–15, 1974.

[62] D.S. Moore and M.C. Spruill. Unified large-sample theory of general chi-squared statistics for tests of fit. *The Annals of Statistics*, 3(3):599–616, 1975.

[63] T. Moreau, J. O'Quigley, and J. Lellouch. On d. schoenfeld's approach for testing the proportional hazards assumption. *Biometrika*, 73(2):513–515, 1986.

[64] T. Moreau, J. O'Quigley, and M. Mesbah. A global goodness-of-fit statistic for the proportional hazards model. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 34(3):212–218, 1985.

[65] G.S. Mudholkar, D.K. Srivastava, and G.D. Kollia. A generalization of the weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, 91(436):1575–1583, 1996.

[66] M. Mu(ller. Goodness-of-fit criteria for survival data. 2004.

[67] N.J.D. Nagelkerke, J. Oosting, and A.A.M. Hart. A simple test for goodness of fit of cox's proportional hazards model. *Biometrics*, 40(2):483–486, 1984.

[68] L. Natarajan, M. Pu, B.A. Parker, C.A. Thomson, B.J. Caan, S.W. Flatt, L. Madlensky, R.A. Hajek, W.K. Al-Delaimy, N. Saquib, et al. Time-varying effects of prognostic factors associated with disease-free survival in breast cancer. *American Journal of Epidemiology*, 169(12):1463–1479, 2009.

[69] J. O'Quigley and F. Pessione. Score tests for homogeneity of regression effect in the proportional hazards model. *Biometrics*, 45(1):135–144, 1989.

[70] F. O'Sullivan. Nonparametric estimation in the cox model. *The Annals of Statistics*, 21(1):124–145, 1993.

[71] M. Parzen and S.R. Lipsitz. A global goodness-of-fit statistic for cox regression models. *Biometrics*, 55(2):580–584, 1999.

[72] L. Peng and Y. Huang. Survival analysis with temporal covariate effects. *Biometrika*, 94(3):719–733, 2007.

[73] M.P. Perme and P.K. Andersen. Checking hazard regression models using pseudo-observations. *Statistics in Medicine*, 27(25):5309–5328, 2008.

[74] L.A. Persson, A. Ivarsson, and O. Hernell. Breast-feeding protects against celiac disease in childhood–epidemiological evidence. *Advances in Experimental Medicine and Biology*, 503:115–123, 2002.

[75] A.N. Pettitt and I.B. Daud. Investigating time dependence in cox's proportional hazards model. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 39(3):313–329, 1990.

[76] C. Quantin, T. Moreau, B. Asselain, J. Maccario, and J. Lellouch. A regression survival model for testing the proportional hazards hypothesis. *Biometrics*, 52(3):874–885, 1996.

[77] P.D. Sasieni and A. Winnett. Martingale difference residuals as a diagnostic tool for the cox model. *Biometrika*, 90(4):899–912, 2003.

[78] W. Sauerbrei, P. Royston, and M. Look. A new proposal for multivariable modelling of time-varying effects in survival data based on fractional polynomial time-transformation. *Biometrical Journal*, 49(3):453–473, 2007.

[79] M. Schemper. Cox analysis of survival data with non-proportional hazard functions. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 41(4):455–465, 1992.

[80] D. Schoenfeld. Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika*, 67(1):145–153, 1980.

[81] D. Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69(1):239–241, 1982.

[82] J. Sin-Ho and S. Wieand. Analysis of goodness-of-fit for cox regression model. *Statistics & Probability Letters*, 41(4):379–382, 1999.

[83] L.A. Sleeper and D.P. Harrington. Regression splines in the cox model with application to covariate effects in liver disease. *Journal of the American Statistical Association*, 85(412):941–949, 1990.

[84] W. Stute and J.L. Wang. The jackknife estimate of a kaplan-meier integral. *Biometrika*, 81(3):602–606, 1994.

[85] P.F. Su, Y. Chi, C.I. Li, Y. Shyr, and Y.D. Liao. Analyzing survival curves at a fixed point in time for paired and clustered right-censored data. *Computational Statistics & Data Analysis*, 55(4):1617–1628, 2011.

[86] T.M. Therneau, P.M. Grambsch, and T.R. Fleming. Martingale-based residuals for survival models. *Biometrika*, 77(1):147–160, 1990.

[87] A.A. Tsiatis. A large sample study of cox's regression model. *The Annals of Statistics*, 9(1):93–108, 1981.

[88] B.W. Turnbull and L. Weiss. A likelihood ratio statistic for testing goodness of fit with randomly censored data. *Biometrics*, 34(3):367–375, 1978.

[89] Z. Valenta and L. Weissfeld. Estimation of the survival function for gray's piecewise-constant time-varying coefficients model. *Statistics in Medicine*, 21(5):717–727, 2002.

[90] P.J.M. Verweij, H.C. van Houwelingen, and T. Stijnen. A goodness-of-fit test for cox's proportional hazards model based on martingale residuals. *Biometrics*, 54:1517–1526, 1998.

[91] L.J. Wei. Testing goodness of fit for proportional hazards model with censored observations. *Journal of the American Statistical Association*, 79(387):649–652, 1984.

[92] A. Winnett and P. Sasieni. Iterated residuals and time-varying covariate effects in cox regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):473–488, 2003.

[93] D.M. Zucker and A.F. Karr. Nonparametric survival analysis with time-dependent covariate effects: A penalized partial likelihood approach. *The Annals of Statistics*, 18(1):329–353, 1990.