

**ESTIMATION OF THE SURVIVAL FUNCTION  
FOR GRAY'S PIECEWISE-CONSTANT  
TIME-VARYING COEFFICIENTS MODEL**

by

Zdeněk Valenta

M.Sc., Charles University, 1988

M.Sc., London School of Hygiene & Tropical Medicine, 1991

M.S., Carnegie Mellon University, 1997

Submitted to the Graduate Faculty of Dept. of Biostatistics,

GSPH, in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

University of Pittsburgh, 2002

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This dissertation was presented by

ZDENĚK VALENTA

It was defended on

April 10, 2002.

and approved by

Stewart J. Anderson, Ph.D., Department of Biostatistics

Derek C. Angus, M.D., M.P.H., Department of Critical Care Medicine

Chung-Chou H. Chang, Ph.D., Department of Medicine

Sati Mazumdar, Ph.D., Department of Biostatistics

Dissertation Director: Lisa A. Weissfeld, Ph.D., Department of Biostatistics

# Abstract

## ESTIMATION OF THE SURVIVAL FUNCTION FOR GRAY'S PIECEWISE-CONSTANT TIME-VARYING COEFFICIENTS MODEL

Zdeněk Valenta, Ph.D.

University of Pittsburgh, 2002

Gray's extension of Cox's proportional hazards (PH) model for right-censored survival data allows for a departure from the PH assumption via introduction of time-varying regression coefficients (TVC) using penalized splines. Gray's work focused on estimation, inference and residual analyses, but no estimator for the survival function has been proposed. We derive a survival function estimator for one important member of the class of TVC models - a piecewise-constant time-varying coefficients (PC-TVC) model. We also derive an estimate for the confidence limits of the survival function. Accuracy in estimating underlying survival times and survival quantiles is assessed for both Cox's and Gray's PC-TVC model using a simulation study featuring scenarios violating the PH assumption. Finally, an example of the estimated survival functions and the corresponding confidence limits derived from Cox's PH and Gray's PC-TVC model, respectively, is presented for a liver transplant data set.

In the second part of the thesis we examine the effect of model misspecification for two

classes of regression models for right-censored survival data - additive and multiplicative models for the conditional hazard rate. A particular attention is given to data exhibiting time-varying regression coefficients. The class of multiplicative models is represented by Cox PH model and Gray's TVC model, respectively, and for additive models we use Aalen's linear model. Both Gray's TVC model and Aalen's linear model incorporate time-varying coefficients. A simulation study is performed to cross-analyze survival data which follows either a multiplicative or an additive model for the conditional hazard rate. The effect of misspecifying the true model for the conditional hazard rate is assessed by looking at the power of the individual models to detect an existing effect, bias and mean square error observed for each conditional model-based estimator of survival. We also show that Aalen's model formulae is a first order Taylor series approximation of that of Gray's model which explains the comparably higher flexibility on part of the Aalen's model as compared to the Cox PH when the Gray's TVC model for the data is misspecified.

# Contents

<b>1</b>	<b>Introduction</b>	<b>10</b>
<b>2</b>	<b>Estimation of the survival function for Gray's piecewise-constant time-varying coefficients model</b>	<b>16</b>
2.1	Abstract . . . . .	16
2.2	Introduction . . . . .	18
2.3	Estimated survival for Gray's PC-TVC model . . . . .	21
2.4	Confidence limits based on the log-transformation . . . . .	24
2.5	Confidence limits based on the log-(log)-transformation . . . . .	26
2.6	Simulation studies . . . . .	28
2.7	UNOS data example . . . . .	33
2.8	Conclusions . . . . .	37
2.9	Acknowledgments . . . . .	37
<b>3</b>	<b>Model misspecification effect in regression models</b>	

<b>for right-censored survival data (with emphasis on the presence of time-varying coefficients)</b>	<b>38</b>
3.1 Abstract . . . . .	38
3.2 Introduction . . . . .	40
3.3 Bias and mean square error of the estimated survival . . . . .	44
3.4 Simulation study design . . . . .	46
3.5 Aalen Data Model . . . . .	49
3.6 Gray Data Model . . . . .	51
3.7 Cox Data Model . . . . .	52
3.8 Summary of the results . . . . .	53
3.8.1 Aalen Data Model with TVC (power trend $\beta(t) = (t + 1)^3$ ) . . . . .	54
3.8.2 Aalen Data Model with TVC absent (constant $\beta(t) = 1$ ) . . . . .	62
3.8.3 Gray Data Model with a linear trend in TVC ( $\beta(t) = 5t$ ) . . . . .	69
3.8.4 Cox Data Model (constant $\beta(t) = 0.5$ ) . . . . .	80
3.9 Conclusions . . . . .	87

3.10 Acknowledgments . . . . . 89

## List of Tables

Table 1: Results Summary for UNOS Cancer Patients (502 observations with 278 failures)

Table 2: True survival probabilities for Aalen Data Model with TVC present (power trend  $\beta(t) = (t + 1)^3$ )

Table 3: Power Analysis for Aalen Data Model with TVC present (power trend  $\beta(t) = (t + 1)^3$ )

Table 4: True survival probabilities for Aalen Data Model with TVC absent (constant  $\beta(t) = 1$ )

Table 5: Power Analysis for Aalen Data Model with TVC absent ( $\beta(t) = 1$ )

Table 6: True survival probabilities for Gray Data Model with a linear trend in TVC ( $\beta(t) = 5t$ )

Table 7: Power Analysis for Gray Data Model with  $\beta(t) = 5t$

Table 8: Power Analysis for Gray Data Model with  $\beta(t) = \log(t)$

Table 9: Power Analysis for Gray Data Model with  $\beta(t) = t - 1$



Table 10: True survival probabilities for Cox Data Model with  $\beta(t) = 0.5$

Table 11: Power Analysis for Cox Data Model with  $\beta(t) = 0.5$

# List of Figures

Figure 1: Simulation studies results summary

Figure 2: UNOS cancer patients post-liver transplant graft survival with 95% C.L.'s,  
Cox and Gray model results for a subject with median-valued covariates

Figure 3: Bias and RMSE for Aalen Data Model with a Power Trend in  $\beta(y)$ , Light  
Censoring

Figure 4: Bias and RMSE for Aalen Data Model with a Power Trend in  $\beta(y)$ , Moderate  
Censoring

Figure 5: Bias and RMSE for Aalen Data Model with a Power Trend in  $\beta(y)$ , Heavy  
Censoring

Figure 6: Bias and RMSE for Aalen Data Model with a Constant Trend in  $\beta(y)$ , Light  
Censoring

Figure 7: Bias and RMSE for Aalen Data Model with a Constant Trend in  $\beta(y)$ ,  
Moderate Censoring

Figure 8: Bias and RMSE for Aalen Data Model with a Constant Trend in  $\beta(y)$ , Heavy  
Censoring

Figure 9: Bias and RMSE for Gray Data Model with a Linear Trend in  $\beta(y)$ , Light Censoring

Figure 10: Bias and RMSE for Gray Data Model with a Linear Trend in  $\beta(y)$ , Moderate Censoring

Figure 11: Bias and RMSE for Gray Data Model with a Linear Trend in  $\beta(y)$ , Heavy Censoring

Figure 12: Bias and RMSE for Cox Data Model with a Constant Trend in  $\beta(y)$ , Light Censoring

Figure 13: Bias and RMSE for Cox Data Model with a Constant Trend in  $\beta(y)$ , Moderate Censoring

Figure 14: Bias and RMSE for Cox Data Model with a Constant Trend in  $\beta(y)$ , Heavy Censoring

# 1 Introduction

In survival analysis the primary function of interest is the hazard rate  $\lambda(t|z)$ , which is closely related to the failure time  $T$ , the dependent variable in survival models. The hazard rate  $\lambda(t|z)$  is defined as a conditional instantaneous failure rate at time  $t$  given a subject's set of covariates  $z$  and his or her survival up to time  $t$ . In this thesis we explore several different types of models that are available for fitting right-censored survival data. The models that we will be dealing with are all concerned with predicting the behavior of the hazard rate over time given the subject's covariate pattern. These models fall into two distinct categories: either additive or multiplicative models for the conditional hazard rate. For the purpose of this thesis, the Cox Proportional Hazards (PH) model <sup>[1]</sup> and Gray's piecewise-constant time-varying coefficients (PC-TVC) model <sup>[2]</sup> will represent the class of multiplicative models, while Aalen's linear model <sup>[3]</sup> will serve as a representative of the class of additive models. The natural distinction between these two classes of models is that in the additive class we model the absolute hazard rate (absolute risk), while in the multiplicative class we model the hazard ratio (relative risk). Another distinction reaching across the two classes of models is based on the inclusion of time varying regression coefficients. Both Aalen's linear model and Gray's PC-TVC model incorporate time-varying coefficients, while the Cox PH model regression coefficients are assumed to remain constant.

An important output from each of these models is the ability to describe the conditional

failure time distribution, or equivalently, the conditional survival distribution of a subject with a covariate pattern  $z$ . The estimated survival function is often of interest when fitting a survival model to data, since this serves as a useful summary of the estimated survival experience of a given population. Gray's work on TVC models has focused on estimation of the model coefficients, inference and residual analysis but no estimator for the survival function has been proposed.

One major contribution of this work is that we have developed a survival function estimator for Gray's PC-TVC model with corresponding estimates of confidence limits based on log- and log(-log)-transformations. The estimation of the survival function is based on the observation that between the successive knots where the hazard regression coefficients are assumed to remain constant, integration with respect to a differential of the cumulative hazard rate may proceed in a manner similar to that for the original Cox PH model. Andersen *et al.* <sup>[4]</sup> (1996) have shown that confidence limits for the survival function estimated from the Cox PH model are optimal when the estimates are based on a log-transformed or log(-log)-transformed scale for the survival curve. Since Gray's PC-TVC model can be viewed as a piecewise PH model, we take a similar approach in developing the variance estimates and corresponding confidence limits on both of these scales.

The Cox PH model has played a prominent role in both the statistical literature and for the analysis of right-censored survival data since its first introduction by Cox in 1972.

The proportional hazards model has been widely used for the analysis of biomedical data from both longitudinal studies and clinical trials. This is mainly due to its appealing mathematical simplicity as well as its general availability through most statistical packages. While the Cox regression model is relatively simple to present, it relies on the assumption of proportionality of hazards which may not be met in all data sets. To address this issue, models that allow for non-proportionality of the conditional hazards through the introduction of time-varying covariate effects have been proposed.

The introduction of the Cox proportional hazards model has resulted in the creation of an extensive body of literature related to the use and further development of this model. Kalbfleisch and Prentice <sup>[5]</sup> (1973) have obtained different marginal likelihoods than those reported by Cox for grouped data, Efron <sup>[6]</sup> (1977) proved the full asymptotic efficiency of Cox's likelihood function for censored data. Peace and Flora <sup>[7]</sup> (1978) considered the size and power of tests related to the regression parameters in the Cox model when the hazard function follows an exponential, Weibull or Gompertz distribution. Lubin <sup>[8]</sup> (1980) studied the analysis of survival data following Cox's failure time model using weighted least squares. Also in 1980, Schoenfeld <sup>[9]</sup> proposed a chi-squared goodness-of-fit test for the proportional hazards regression model by comparing the observed and expected numbers of events in cells arising from a partition of the Cartesian product of the range of covariates and the time axis. A key criticism of this approach is based on the arbitrariness of the partitioning. Two years later Schoenfeld <sup>[10]</sup> introduced

partial residuals for the proportional hazards regression model. The first successful approach to obtaining the asymptotic properties of Cox's estimates is given in Tsiatis<sup>[11]</sup> (1981). Asymptotic theory for Cox's regression model in the framework of counting processes has been introduced by Andersen and Gill<sup>[12]</sup> in 1982. Their study enabled the construction of confidence intervals for the survival function estimated from the Cox model. Also in 1982, Johnson *et al.*<sup>[13]</sup> published a small-sample covariate analysis study of survival data following Cox's model, and Stewart and Pierce<sup>[14]</sup> assessed the efficiency of Cox's model in estimating regression parameters relative to parametric alternatives for grouped data settings. Bailey<sup>[15]</sup> (1983) studied the asymptotic joint distribution of the regression and survival parameter estimates in the Cox regression model. In 1984, Bailey<sup>[16]</sup> showed that the Cox estimator and the general ML estimator of the regression and survival parameters in the Cox model are asymptotically equivalent. Gore *et al.*<sup>[17]</sup> (1984) examined the performance of a wide variety of parametric survival models and Cox's regression model in the context of non-proportional hazards, and suggested a step-function proportional hazards approach as an attractive alternative to the Cox model for this setting. Link<sup>[18]</sup> (1984) studied confidence intervals for the survival function using Cox's proportional hazards regression model with censored data. A martingale approach to understanding Cox's regression model has been further investigated by Gill<sup>[19]</sup> also in 1984.

One of the more useful spline-based extensions of the Cox proportional hazards model

is that proposed by Gray <sup>[2]</sup>. Gray's TVC extension of the Cox PH model employs products of the covariates of interest with the spline functions of time. This allows for a flexible approach to the modeling of covariate effects without necessarily adhering to the assumption of proportional hazards, which may not be satisfied. The most appealing model within the framework of models proposed by Gray is the piecewise-constant TVC (Gray's PC-TVC) model since this model is similar to the original Cox PH model and retains much of the mathematical simplicity of the Cox model. The advantage of the PC-TVC models is their flexibility, since the proportional hazards assumption is only required for each of the time intervals between the successive knots (i.e. time points allowing for a change in the regression coefficients). Gray's PC-TVC model may therefore be viewed as a piecewise proportional hazards model for the conditional hazard rate. A family of models which can be used to model nonproportional data, the time-varying coefficient (TVC) models, have been considered already in 1987 by Gamerman and West <sup>[20]</sup>, and further developed by Zucker and Karr <sup>[21]</sup> in 1990. Building on the work of Tsiatis <sup>[11]</sup>, Andersen and Gill <sup>[12]</sup> and Gill <sup>[19]</sup>, O'Sullivan <sup>[22]</sup> (1993) treated nonparametric estimation in the Cox model using an approach complementary to that of Zucker and Karr <sup>[21]</sup>. The methodology of Zucker and Karr was further developed by Gray <sup>[2, 23]</sup>. Time-varying coefficient models were also studied by Hastie and Tibshirani <sup>[24]</sup> (1993) and the use of regression splines in modeling the conditional hazard rate is discussed in Sleeper and Harrington <sup>[25]</sup> (1990) and Gray <sup>[2]</sup>. The use of time dependence in Cox's PH model was also investigated by Pettit and Daud <sup>[26]</sup>.



In the second part of this thesis we examine the effect of model misspecification on regression models for right-censored survival data involving a single continuous covariate. We perform a simulation study to cross-analyze survival data that follows either Aalen's, Cox's PH or Gray's TVC model for the conditional hazard rate. Particular attention is given to data from a model with time-varying regression coefficients. It is often the case that the right-censored survival data are analyzed without prior verification of the assumptions of an assumed model. The effect of misspecifying the true model is assessed by looking at the power of Aalen's, Cox's and Gray's model, respectively, to detect an existing effect and by comparing the mean square error and bias of the conditional estimator of survival obtained for each of the three models used in analyzing the data. Formulas for bias and mean square error of the conditional survival distribution rely on the ability of the fitted model to estimate survival. In this part we see a direct application of the concepts developed in the first part of the thesis. Specifically, the survival function estimator for Gray's PC-TVC model <sup>[27]</sup> (Valenta and Weissfeld, 2002) we developed earlier is now being used for the evaluation of the bias, mean square error and the true survival function when Gray's model is assumed. We also show that Aalen's model formulae appears to be a first order Taylor series approximation of that of Gray's TVC model. This is also supported by our empirical finding that when the true underlying model for the data is multiplicative with TVC present, Aalen's linear model shows a comparably higher flexibility under misspecification than Cox PH model.

## 2 Estimation of the survival function for Gray's piecewise-constant time-varying coefficients model

### 2.1 Abstract

Gray's extension of Cox's proportional hazards (PH) model for right-censored survival data allows for a departure from the PH assumption via introduction of time-varying regression coefficients (TVC). For this model estimation of the conditional hazard rate relies on the inclusion of penalized splines. Cubic penalized splines tend to be unstable in the right tail of the distribution and thus quadratic, linear and piecewise-constant penalized splines may be a favorable choice. Gray's work focused on estimation, inference and residual analyses, but no estimator for the survival function has been proposed.

We derive a survival function estimator for one important member of the class of TVC models - a piecewise-constant time-varying coefficients (PC-TVC) model. Using the first-order Taylor series approximation we also derive an estimate for the variance of the log- and log(-log)-transformed survival function, which in turn leads to estimated confidence limits on the corresponding scales of the survival function. Accuracy in estimating underlying survival times and survival quantiles is assessed for both Cox's and Gray's PC-TVC model using a simulation study featuring scenarios violating the PH assumption.

Finally, an example of the estimated survival functions and the corresponding confidence limits derived from Cox's PH and Gray's PC-TVC model, respectively, is presented for a liver transplant data set.

## 2.2 Introduction

The Cox proportional hazards (PH) model has played a prominent role in both the statistical literature and for the analysis of right-censored survival data since its first introduction by Cox [1] in 1972. It has been widely used for the analyses of biomedical data from both longitudinal studies and clinical trials, mainly due to its appealing mathematical simplicity, as well as its general availability through most statistical packages.

While the Cox PH model is relatively simple to present, it relies on the assumption of proportionality which may not be met in all data sets. To address this issue, models that allow for non-proportionality of the conditional hazards through the introduction of penalized splines have been proposed. A family of models which can be used to model non-proportional data, the time-varying coefficient (TVC) models, have been considered by Gamerman and West [20], and Zucker and Karr [21]. A general treatment of the first order asymptotic analysis of the penalized likelihood is due to Cox and O’Sullivan [28]. Building on the work of Tsiatis [11], Andersen and Gill [12] and Gill [19], O’Sullivan [22] treated nonparametric estimation in the Cox model using an approach complementary to that of Zucker and Karr [21].

The methodology of Zucker and Karr was further developed by Gray [2, 23]. Time-varying coefficient models were also studied by Hastie and Tibshirani [24] and the use of regression splines in modeling the conditional hazard rate is discussed in Sleeper and Harrington [25] and Gray [2]. The use of time dependence in Cox’s PH model was also

investigated by Pettitt and Daud [26], Hess [29] and Verweij and van Houwelingen [30]. One of the more useful spline-based extensions of the Cox proportional hazards model is that proposed by Gray [2]. Gray's TVC extension of the Cox PH model employs products of the covariates of interest with the spline functions of time. This allows for a flexible approach to the modeling of covariate effects without necessarily adhering to the assumption of proportional hazards, which may not be satisfied.

The most appealing model within the framework of models proposed by Gray is the piecewise-constant TVC (Gray's PC-TVC) model since this model is similar to the original Cox PH model and retains much of the mathematical simplicity of the Cox model. The advantage of the PC-TVC models is their flexibility, since the proportional hazards assumption is only required for each of the time intervals between the successive knots (i.e. time points allowing for a change in the regression coefficients). Gray's PC-TVC model may therefore be viewed as a piecewise proportional hazards model for the conditional hazard rate. The estimated survival function is often of interest when fitting a survival model to data, since this serves as a useful summary of the estimated survival experience of a given population. Gray's work on TVC models has focused on estimation of the model coefficients, inference and residual analysis and to date, no estimator for the survival function has been presented.

Andersen *et al.* [4] show that confidence limits for the survival function estimated from the Cox PH model are optimal when the estimates are based on a log-transformed or

log(-log)-transformed scale for the survival curve. In this paper we present an estimator of the survival function under Gray's PC-TVC model. Estimation is based on the observation that between the successive knots, where the hazard regression coefficients are assumed to remain constant, the integration with respect to a differential of the cumulative hazard rate may proceed in a manner similar to that for the original Cox PH model. The estimated variance of the predicted survival function under Gray's PC-TVC model is derived for both the log- and log(-log)-transformed scale of the survival function and corresponding estimates of the confidence limits are presented.

### 2.3 Estimated survival for Gray's PC-TVC model

Within the TVC family of models we assume that the hazard function can be modeled as follows:

$$d\Lambda(t|x) = d\Lambda_0(t) \exp\{x'\beta(t)\}, \quad (1)$$

where  $\Lambda(\cdot)$  denotes the cumulative hazard function and  $\Lambda_0(\cdot)$  denotes the cumulative baseline hazard. Here  $\beta'(t) = (\beta_1(t), \beta_2(t), \dots, \beta_p(t))$ , where  $\beta_j(t) = \sum_k \theta_{jk} B_{jk}(t)$ ,  $j = 1, \dots, p$  [2] are modeled with a full set of B-spline basis functions,  $B_{jk}(t)$  [31]. Unlike Cox's proportional hazards model where the hazard regression coefficients,  $\beta(t)$ , in (1) are fixed, they are a function of time under Gray's PC-TVC model. Specifically, the coefficients are assumed to be constant only for values of  $t \in [\tau_j, \tau_{j+1})$ ,  $j = 0, \dots, q$ . Here  $\tau_j$ ,  $j = 1, \dots, q$ , denote the internal knots,  $\tau_0 = 0$ , and  $\tau_{q+1} = T$  represent the maximum observed (survival or censoring) time. Under Gray's PC-TVC model, the coefficients,  $\beta(t)$ , are therefore right-continuous step functions of time with jumps possibly occurring at the knots  $\tau_j$ ,  $j = 1, \dots, q$ .

Estimation of the regression parameters in Gray's PC-TVC model proceeds by maximizing the penalized partial likelihood, which involves a partial likelihood term as in the Cox model, plus the following penalty term:  $\frac{1}{2} \lambda_j \sum_{k=2}^{q+1} (\theta_{jk} - \theta_{j,k-1})^2$ , where  $q$  is the number of internal knots for modeling the splines [2]. An essential component of the survival function estimate under Gray's PC-TVC model is based on the corresponding estimate of the cumulative baseline hazard. We extend Breslow's estimator [32] of the

cumulative baseline hazard function to derive an estimator of the baseline hazard function for the TVC model. We assume that the coefficients,  $\beta$ , in Breslow's formula can simply be replaced with their corresponding time-varying counterparts,  $\beta(t)$ .

Consequently, under the TVC model (1) for the conditional hazard rate the estimated cumulative baseline hazard function is of the form:

$$\hat{\Lambda}_0(t) = \int_0^t \frac{1}{\sum_i Y_i(s) \exp\{z'_i \hat{\beta}(s)\}} \sum_{i=1}^n dN_i(s) \quad , \quad (2)$$

where  $Y_i(t)$  is an indicator function for the  $i$ -th patient's risk status at time  $t$  (i.e.,  $Y_i(t) = 1$  if the  $i$ -th patient is in the risk set at time  $t$ , and 0 otherwise).

For Gray's PC-TVC model the formula for the estimated survival function of a patient with  $p$ -variate covariate vector,  $z_0$ , will be:

$$\begin{aligned} \hat{S}(t|z_0) &= \exp \left\{ - \int_0^t d\hat{\Lambda}(s|z_0) \right\} = \exp \left\{ - \int_0^t \exp\{z'_0 \hat{\beta}(s)\} d\hat{\Lambda}_0(s) \right\} \\ &= \exp \left\{ - \int_0^T I(s \leq t) \exp\{z'_0 \hat{\beta}(s)\} d\hat{\Lambda}_0(s) \right\}. \end{aligned} \quad (3)$$

On the log-transformed scale of the survival function we obtain:

$$\log \hat{S}(t|z_0) = - \int_0^T I(s \leq t) \exp\{z'_0 \hat{\beta}(s)\} d\hat{\Lambda}_0(s) = - \sum_{j=0}^q \exp\{z'_0 \hat{\beta}(\tau_j)\} \hat{\Lambda}_{0j}(t), \quad (4)$$

where

$$\hat{\Lambda}_{0j}(t) = \int_{[\tau_j, \tau_{j+1})} I(s \leq t) d\hat{\Lambda}_0(s) = \int_{[\tau_j, \tau_{j+1})} I(s \leq t) \frac{\sum_{i=1}^n dN_i(s)}{\sum_i Y_i(s) \exp\{z'_i \hat{\beta}(s)\}} \quad (5)$$

represents a contribution to the estimated (total) cumulative baseline hazard  $\hat{\Lambda}_0(t)$  corresponding to an interval  $[\tau_j, \tau_{j+1})$ . Since  $\beta(\tau_j)$  remains constant on  $[\tau_j, \tau_{j+1})$ , we will



make use of the following notation:  $\beta_j = \beta(\tau_j)$ , where  $\beta_j$  is a vector of length  $p$ . Given a covariate vector  $z_0$ , we thus obtain an estimate of the survival function,  $S(t|z_0)$ , as follows:

$$\hat{S}(t|z_0) = \exp \left\{ - \sum_{j=0}^q \exp \{ z_0' \hat{\beta}_j \} \hat{\Lambda}_{0j}(t) \right\}. \quad (6)$$

## 2.4 Confidence limits based on the log-transformation

Based on (4), the formula for the variance of the log-transformed estimator of the survival function is as follows:

$$\begin{aligned} \text{Var} \left( \log \hat{S}(t|z_0) \right) &= \text{Cov} \left( - \sum_{j=0}^q \hat{\Lambda}_{0j}(t) \exp \left( z'_0 \hat{\beta}_j \right), - \sum_{j=0}^q \hat{\Lambda}_{0j}(t) \exp \left( z'_0 \hat{\beta}_j \right) \right) = \\ &= \sum_{k=0}^q \sum_{l=0}^q \text{Cov} \left( \hat{\Lambda}_{0k}(t) \exp \left( z'_0 \hat{\beta}_k \right), \hat{\Lambda}_{0l}(t) \exp \left( z'_0 \hat{\beta}_l \right) \right) . \end{aligned} \quad (7)$$

Note that (7) requires an estimator of the covariance which can be derived from a Taylor series approximation.

We also define the following functions:

$$g(\hat{\beta}_j, t) = \hat{\Lambda}_{0j}(t) \exp \left( z'_0 \hat{\beta}_j \right), \quad j \in \{0, \dots, q\}. \quad (8)$$

The vector of the corresponding partial derivatives may be evaluated as follows:

$$\frac{\partial}{\partial \hat{\beta}_j} g(\hat{\beta}_j, t) = \exp \left( z'_0 \hat{\beta}_j \right) \left( z_0 \hat{\Lambda}_{0j}(t) + \frac{\partial}{\partial \hat{\beta}_j} \left( \hat{\Lambda}_{0j}(t) \right) \right).$$

Now, the first order Taylor series approximation of  $g(\hat{\beta}_j)$  about the expected value of  $\hat{\beta}_j$  (which we will denote by  $\beta_j$ ) can be written as:

$$g(\hat{\beta}_j, t) \approx g(\beta_j, t) + \left[ \frac{\partial}{\partial \hat{\beta}_j} \left( g(\hat{\beta}_j, t) \right) \Big|_{\hat{\beta}_j = \beta_j} \right]' \left( \hat{\beta}_j - \beta_j \right) . \quad (9)$$

The covariance terms in (7) can be approximated at time  $t$  using the Delta method as follows:

$$\text{Cov} \left\{ g(\hat{\beta}_k, t), g(\hat{\beta}_l, t) \right\} \approx W_k(t)' \text{Cov} \left( \hat{\beta}_k, \hat{\beta}_l \right) W_l(t), \quad (10)$$

where

$$W_j(t) = \left[ \exp(z_0' \hat{\beta}_j) \left( z_0 \hat{\Lambda}_{0j}(t) + \frac{\partial}{\partial \hat{\beta}_j} \hat{\Lambda}_{0j}(t) \right) \right]_{|\hat{\beta}_j = \beta_j}, \quad j \in \{k, l\} \quad (11)$$

and

$$\frac{\partial}{\partial \hat{\beta}_j} \hat{\Lambda}_{0j}(t) = \int_{[\tau_j, \tau_{j+1})} I(s \leq t) \frac{-\sum_i Y_i(s) z_i \exp\{z_i' \hat{\beta}_j\}}{\left\{ \sum_i Y_i(s) \exp\{z_i' \hat{\beta}_j\} \right\}^2} \sum_{i=1}^n dN_i(s), \quad j \in \{k, l\} \quad (12)$$

is a p-variate vector of partial derivatives of  $\hat{\Lambda}_{0j}(t)$ .

At time t we also have:

$$z_0 \hat{\Lambda}_{0j}(t) = \int_{[\tau_j, \tau_{j+1})} I(s \leq t) \frac{\sum_i Y_i(s) z_0 \exp\{z_i' \hat{\beta}_j\}}{\left\{ \sum_i Y_i(s) \exp\{z_i' \hat{\beta}_j\} \right\}^2} \sum_{i=1}^n dN_i(s), \quad (13)$$

so that

$$W_j(t) = \left[ \int_{[\tau_j, \tau_{j+1})} I(s \leq t) \frac{\sum_i Y_i(s) (z_0 - z_i) \exp\{(z_0 + z_i)' \hat{\beta}_j\}}{\left\{ \sum_i Y_i(s) \exp\{z_i' \hat{\beta}_j\} \right\}^2} \sum_{i=1}^n dN_i(s) \right]_{|\hat{\beta}_j = \beta_j}. \quad (14)$$

Consequently, the formula for the estimated variance of the predicted survival function will take the following form:

$$\hat{\text{Var}}(\log \hat{S}(t|z_0)) = \sum_{k=0}^q \sum_{l=0}^q W_k(t)' \text{Cov}(\hat{\beta}_k, \hat{\beta}_l) W_l(t). \quad (15)$$

Finally, the 100(1- $\alpha$ )% Confidence Limits for the survival function estimated under Gray's PC-TVC model are calculated as follows:

$$\exp \left( \log \hat{S}(t|z_0) \pm z_{1-\alpha/2} \sqrt{\hat{\text{Var}}(\log \hat{S}(t|z_0))} \right), \quad (16)$$

where  $z_{1-\alpha/2}$  denotes an upper  $\alpha/2$ -quantile of the standard normal distribution,  $\hat{\text{Var}}(\log \hat{S}(t|z_0))$

is given by (15) and  $\log \hat{S}(t|z_0)$  is estimated based on equation (4).

## 2.5 Confidence limits based on the log-(log)-transformation

On the log(-log)-scale of the estimated survival function we obtain the following:

$$\log\left(-\log\left(\hat{S}(t|z_0)\right)\right) = \log\left(\sum_{j=0}^q \hat{\Lambda}_{0j}(t) \exp\left(z'_0 \hat{\beta}_j\right)\right). \quad (17)$$

Let us denote the complete vector of time-varying coefficient estimates from Gray's PC-TVC model by  $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_q)$ . Note that each component of the vector is itself a vector of length p (where p stands for the number of covariates being modeled by splines).

Also, let  $\frac{\partial}{\partial \hat{\beta}} \tilde{g}(\hat{\beta}) = \left(\frac{\partial}{\partial \hat{\beta}_0} \tilde{g}(\hat{\beta}), \frac{\partial}{\partial \hat{\beta}_1} \tilde{g}(\hat{\beta}), \dots, \frac{\partial}{\partial \hat{\beta}_q} \tilde{g}(\hat{\beta})\right)$ , where each of the q components of the vector of partial derivatives of  $\tilde{g}(\hat{\beta})$  is itself a vector of length p.

Using this notation we write at time t:

$$\tilde{g}(\hat{\beta}, t) = \log\left(\sum_{j=0}^q \hat{\Lambda}_{0j}(t) \exp\left(z'_0 \hat{\beta}_j\right)\right). \quad (18)$$

Thus the k-th component of the vector of partial derivatives (being itself a vector of length p) will be:

$$\left(\frac{\partial}{\partial \hat{\beta}} \tilde{g}(\hat{\beta}, t)\right)_k = \frac{\exp\left(z'_0 \hat{\beta}_k\right) \left(z_0 \hat{\Lambda}_{0k}(t) + \frac{\partial}{\partial \hat{\beta}_k} (\hat{\Lambda}_{0k}(t))\right)}{\sum_{j=0}^q \exp\left(z'_0 \hat{\beta}_j\right) \hat{\Lambda}_{0j}(t)}. \quad (19)$$

It follows from (14) that:

$$\left(\frac{\partial}{\partial \hat{\beta}} \tilde{g}(\hat{\beta}, t)\right)_k = \int_{[\tau_k, \tau_{k+1})} I(s \leq t) \frac{\sum_i Y_i(s) (z_0 - z_i) \exp\{(z_0 + z_i)' \hat{\beta}_k\}}{\left\{\sum_{j=0}^q \exp\{z'_0 \hat{\beta}_j\} \hat{\Lambda}_{0j}(t)\right\} \left\{\sum_i Y_i(s) \exp\{z'_i \hat{\beta}_k\}\right\}^2} \sum_{i=1}^n dN_i(s). \quad (20)$$

Let us write:

$$\tilde{W}(t) = \left[ \frac{\partial}{\partial \beta} \tilde{g}(\hat{\beta}, t) \right]_{|\hat{\beta}=\beta}. \quad (21)$$

Using the first-order Taylor series approximation of the log(-log)-transformed survival function we can estimate the variance as follows:

$$\hat{\text{Var}} \left( \log \left( -\log \left( \hat{S}(t|z_0) \right) \right) \right) \approx \tilde{W}(t)' \text{Var}(\hat{\beta}) \tilde{W}(t), \quad (22)$$

where  $\text{Var}(\hat{\beta})$  is the covariance matrix of the complete vector of time-varying coefficients with the partial derivatives in expression (22) evaluated as in (20).

Consequently, the  $100(1-\alpha)\%$  confidence limits for the survival function estimated under Gray's PC-TVC model based on the log(-log) transformation of the survival function will be given by:

$$\exp \left\{ -\exp \left\{ \log \left( -\log \hat{S}(t|z_0) \right) \mp z_{1-\alpha/2} \sqrt{\hat{\text{Var}} \left( \log \left( -\log \hat{S}(t|z_0) \right) \right)} \right\} \right\}, \quad (23)$$

where  $z_{1-\alpha/2}$  denotes an upper  $\alpha/2$ -quantile of the standard normal distribution,  $\log \hat{S}(t|z_0)$  is obtained from (4) and  $\hat{\text{Var}} \left( \log \left( -\log \hat{S}(t|z_0) \right) \right)$  is estimated using (22).

## 2.6 Simulation studies

In order to assess the accuracy of both Cox's and Gray's survival estimators we designed two simulation studies allowing for comparison of the estimated survival quantiles and probabilities of survival obtained from Cox's and Gray's model with the true underlying values. We considered scenarios that violate the assumption of proportionality. In all instances throughout this article, Gray's PC-TVC model was fitted with 10 knots selected automatically so that approximately the same number of events was observed between the successive knots, and 4 degrees of freedom that fully specify the choice of the corresponding value of the smoothing parameter.

We have generated survival data from the piecewise-exponential distribution with two time-points allowing for a change in the hazard at .3 and .8 years. For a set of survival probabilities  $\{.99., .95, .90, .75, .50, .25, .10, .05, .01\}$ , the corresponding time-points were estimated using both Cox's and Gray's models based on 1000 samples of size 150. Also, for a set of time points of 3, 7, 14 and 30 days and .5, 1, 1.5 and 3 years, estimates of the corresponding probabilities of survival were calculated from each of the models. For this simulation study, all of the data are complete. The results we obtained for censored data were very similar to those for complete data. The introduction of censoring, however, leaves some quantities related to the right tail of the distribution inestimable (e.g. time points corresponding to small survival probabilities).

In the first study, one third of each sample (associated with the first covariate being

an indicator function for that group) was generated with hazards of  $(1.5, 1, 2)$ , the second third of the sample (associated with the second covariate) was generated with the hazards reversed (i.e.  $(2, 1, 1.5)$ ), and the baseline hazards were all taken from an exponential distribution with parameter 1. In the second study, hazards of  $(2, 1, .5)$  were associated with the first covariate, those reversed  $(.5, 1, 2)$  were associated with the second covariate and a constant hazard of 1 was again assumed for the baseline.

We wrote two simple Splus functions to compute the true survival quantiles and probabilities for the piecewise-exponential distribution. Figure 1 (consisting of 4 panels numbered clockwise beginning from the top left quadrant) presents plots of the differences between the estimated and the true quantities (i.e. probabilities and survival quantiles respectively), as determined in both of the above studies. In both studies the survival curves were estimated at the covariate values  $(1,0)$  and  $(0,1)$  respectively, indicating a patient exhibiting the hazards specified by the first or second (i.e. reversed) set of hazards used in each example.

Panels 1 and 4 of Figure 1, based on 1000 samples, reveal that the differences between the true and estimated survival quantiles (times) were consistently smaller for Gray's model (denoted by circles in the plot). For this model the four corresponding trends in the hazard implied average departures from the true underlying quantiles of less than 20 days with the exception of the 1% quantile, for which the average departures ranged from 21 to 60 days. For the Cox model (denoted by triangles in the plot), however,

departures from the true values greater than 50 days were observed for the 75, 50, 10, 5 and 1% survival quantiles. Panel 4 reveals that the estimates of the two smallest survival quantiles based on the Cox model were actually off by more than 1 year for both trends in the hazard. The magnitude of error observed was generally higher for the hazard rates of (2,1,.5) or reversed, than for those of (1.5,1,2) or reversed.

Similarly, panels 2 and 3 of Figure 1 illustrate the superior performance of Gray's PC-TVC model over that of the Cox model in terms of the accuracy of the estimated survival probabilities associated with several pre-determined time points. For 1000 samples simulated with hazards (1.5,1,2) and (2,1,1.5) respectively, the probability estimates based on Gray's model were all within a distance of .01 from the true underlying values. For hazard rates of (2,1,.5) and (.5,1,2), estimates obtained from Gray's model exceeded the .01 distance in 3 of 18 cases with the maximum departure from the true value being .017 (associated with the time point of 6 months). Based on the Cox model, however, departures below .01 were observed in only 10 of 36 cases. In 16 of the 36 cases the magnitude of error associated with the Cox model exceeded the level of .025. The magnitude of error was again generally higher for the hazard rates (2,1,.5) or reversed, than for those of (1.5,1,2) or reversed.

The averaging effect of the Cox model is well documented in panels 2 and 3. Since the simulated hazard rates stabilized after .8 years, we observe that the departures from the true underlying values decreased dramatically after 1 year. As a result of the lack of



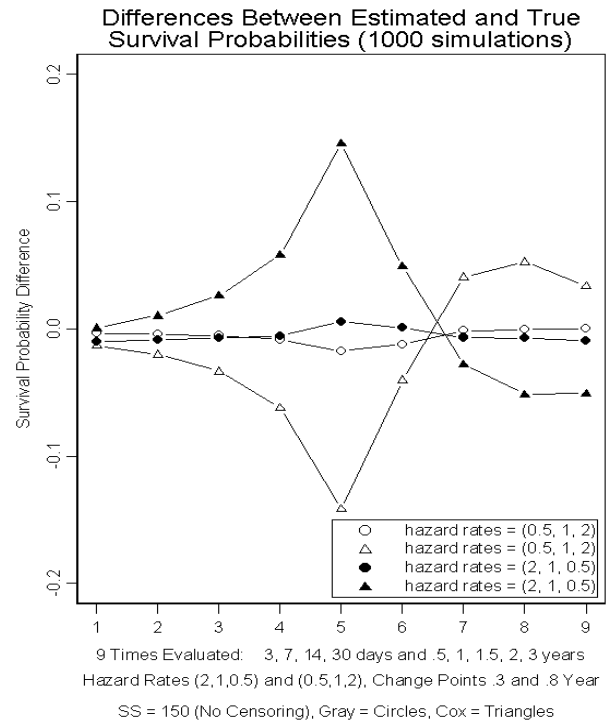
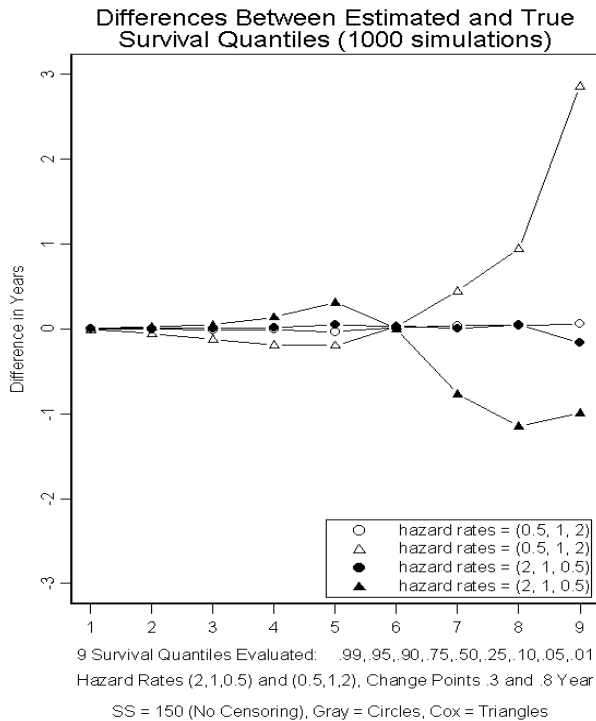
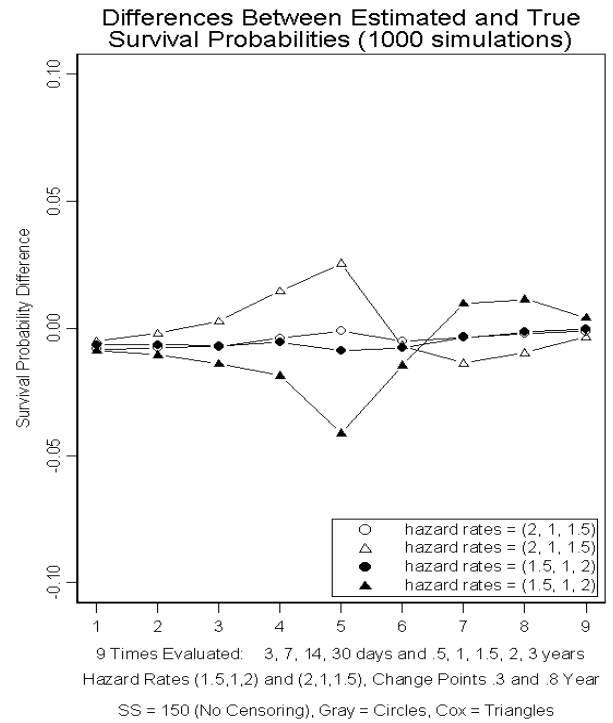
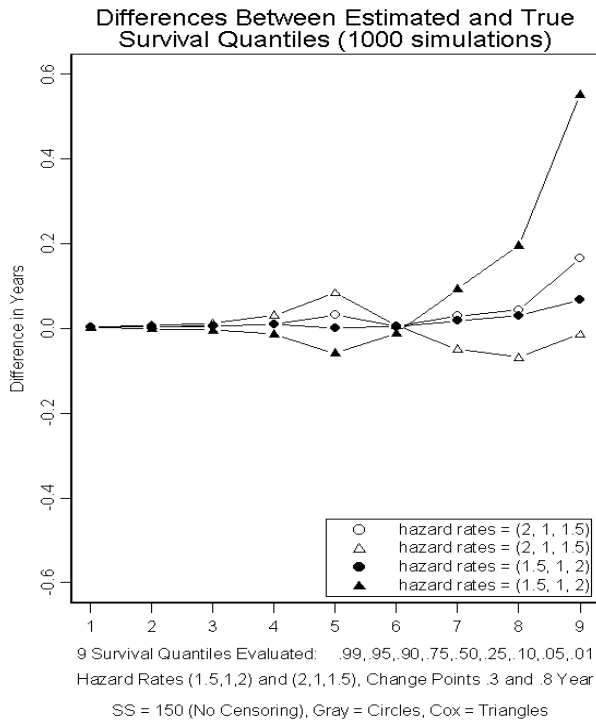


Figure 1: Simulation studies results summary

flexibility on the part of Cox's model, however, this lead to subsequent departures in the opposite directions at the right tail of the distribution.

Results obtained from the simulation studies indicate that a high level of accuracy is maintained by the survival function estimates based on Gray's model, even in the tails of the distribution. Estimates obtained using Gray's model were generally close to the true values, while those derived from the Cox model occasionally showed large departures from the true underlying values. This resulted from a violation of the proportionality assumption in the data. The lack of precision in Cox's model was caused by the averaging of the time-varying effects, which is a built-in feature of Cox's model. In contrast, a high level of accuracy has been maintained by Gray's survival estimator, even in the tails of the distribution.

## 2.7 UNOS data example

In this section we present a real data example comparing survival function estimators derived from Cox's and Gray's model, respectively. It features a dataset from the UNOS (United Network for Organ Sharing) database of cancer patients who underwent a liver transplant. Here we estimate the graft survival for a subject whose covariate values are set to the median sample values. In graft survival analysis a failure is defined as an organ failure or a death of the recipient. We compare the best Cox and Gray models found for the data. The best models featured the following covariates (with corresponding sample median values listed in the parentheses): donor's anti CMV IGG result (dcmvgr, 1), indicator of whether the recipient had any prior transplant (priortx, 0), log-serum creatinine (lcreat, 0), log-total serum bilirubin (ltbili, 1.224), blood match indicator (abo.mtch, 1) and log-prothrombin time (lptp, 2.695). A summary of the modeling results may be found in Table 1.

**TABLE 1: Results Summary for UNOS Cancer Patients****(502 observations with 278 failures)**

Covariates	Cox's Model			Gray's Model		
	Coeff	p-value	n.prop.	Coeff (Range)	p-value	n.prop.
lcreat	0.266	0.014	0.789	(0.154:0.555)	0.001	0.277
ltbili	0.182	0.001	0.050	-	-	-
dcmvgr	0.307	0.011	0.936	-	-	-
abo.mtch	-1.147	0.049	0.642	(-2.936:0.205)	0.007	0.142
lptp	-	-	-	(-0.244:1.688)	0.000	0.007
priortx	-	-	-	(-5.999:3.228)	0.040	0.769

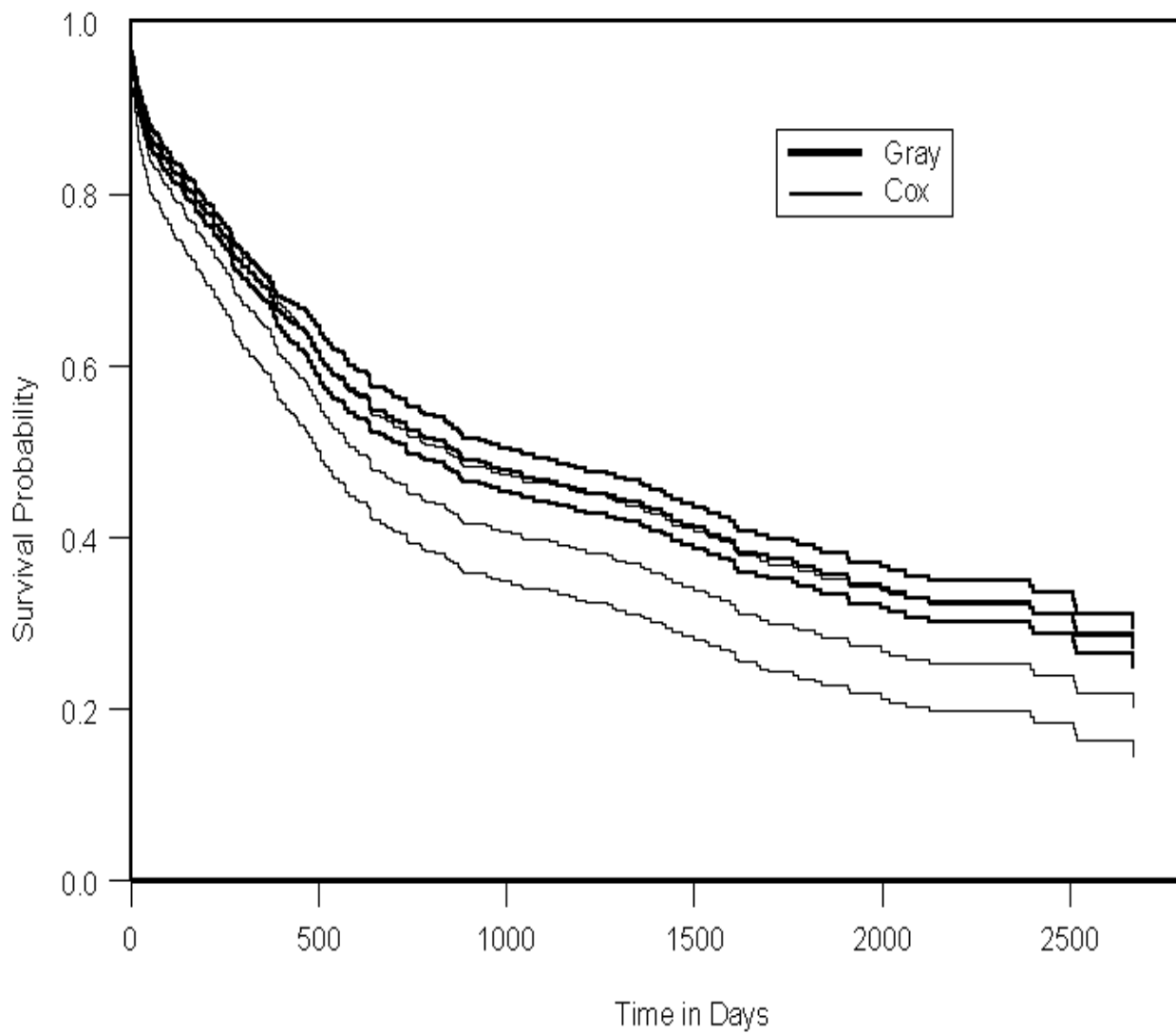


Figure 2: UNOS cancer patients post-liver transplant graft survival with 95% C.L.'s, Cox and Gray model results for a subject with median-valued covariates

Covariates found to be significant under the best Cox model for the liver transplant graft survival of UNOS cancer patients were “lcreat”, “ltbili”, “dcmvgr” and “abo.mtch”, with log-total serum bilirubin (ltbili) being identified as marginally non-proportional with regard to the effect on the hazard rate (p-value 0.0499). The best Gray’s model included “lcreat”, “lptp”, “abo.mtch” and “priortx”. Here the log-prothrombin time (lptp) was identified as having a highly non-proportional effect on the hazard rate (p-value 0.007). Survival functions and 95% confidence limits estimated by the two models at the sample median covariate values are presented in Figure 2. Although the confidence bands for the two survival curves overlap (Gray’s estimated survival function actually follows closely the upper confidence band estimated by the Cox’s model), we can still observe a notable difference between the two survival estimates.

The real data example of this section further illustrates the differences in survival estimates that might be obtained for data which does not follow the proportional hazards assumption.

## 2.8 Conclusions

Gray's piecewise-constant time-varying coefficients model for right-censored survival data is a flexible alternative to the Cox proportional hazards model in scenarios where the PH assumption may not be satisfied. The survival function estimator that we derived for this model provides a useful summary of the modeling results based on the patient's covariate values.

Simulation studies presented earlier have shown a lack of accuracy on the part of the Cox model with regard to estimating survival probabilities and predicting survival quantiles when the survival distribution does not satisfy the PH assumption.

Finally, based on Cox and Gray's model, respectively, a differing graft survival experience was demonstrated for a UNOS cancer patient after a liver transplant.

## 2.9 Acknowledgments

This work was supported in part by the Agency for Health Care Research and Quality (grant R01 HS09694-03) and by the Department of the Army (grant DAMD17-99-1-9536). We thank Dr. Robert J. Gray for making his own implementation of the modeling routine available to us. The survival function estimator has been written as an Splus function based on Gray's routine and is available from the authors upon request.

# **3 Model misspecification effect in regression models for right-censored survival data (with emphasis on the presence of time-varying coefficients)**

## **3.1 Abstract**

We examine the effect of model misspecification in univariate regression models for right-censored survival data. A simulation study is performed to cross-analyze survival data which follows either a multiplicative or an additive model for the conditional hazard rate. We consider two different multiplicative models, the proportional hazards (PH) model of Cox (1972) and Gray's piecewise-constant time-varying coefficients (PC-TVC) model (1992). For the additive model, we use Aalen's linear model (1989) which also incorporates time-varying coefficients. Particular attention is given to data exhibiting time-varying regression coefficients.

The effect of misspecifying the true model for the conditional hazard rate is assessed by looking at the power of the individual models to detect an existing effect and by the observed bias and mean square error for each conditional model-based estimator of survival. We also show that Aalen's model formulae appears to be a first order Taylor series approximation of that of Gray's TVC model. This is also supported by our empirical finding that when the true underlying model for the data is multiplicative



with TVC present, Aalen's linear model shows a comparably higher flexibility under misspecification than Cox PH model.

## 3.2 Introduction

The body of existing literature examining the effect of model misspecification for right-censored survival data is not very extensive. Based on the theoretical findings of Cox and Reid <sup>[33]</sup> (1984) regarding parameter orthogonality and approximate conditional inference, Solomon <sup>[34]</sup> examined the problem of misspecifying regression models in the analysis of survival data by considering accelerated life (AL) and proportional hazards (PH) models as the two alternatives. In the case of independent censoring, she found that to first order regression parameters were proportional under both alternative (i.e. AL and PH) models. This issue was revisited by Hutton and Solomon <sup>[35]</sup> in 1997, when they proposed a one-parameter family mixture survival model which includes both the accelerated life and proportional hazards models. Struthers and Kalbfleisch <sup>[36]</sup> (1986) also examined the problem of misspecifying the Cox PH model when the true model for the data was an accelerated failure time model or when some covariates were omitted from the model. Inference for misspecified parametric survival models was discussed by Hjort <sup>[37]</sup> in 1992. The effects of model misspecification in estimating covariate effects and robustness of the Cox PH model when the true model for the data took the form of several parametric models were studied by Li, Klein and Moeschberger <sup>[38]</sup> in 1996.

In the second part of this thesis we examine the effect of model misspecification for two classes of non-parametric survival models frequently used for the analysis of right-censored survival data, additive and multiplicative models for the conditional hazard

rate. The additive class will be represented by Aalen's linear model <sup>[3]</sup>, while the multiplicative class will be represented by the Cox proportional hazards model <sup>[1]</sup> and Gray's piecewise-constant time-varying coefficients model <sup>[2]</sup>. Both Aalen's linear model and Gray's TVC model incorporate time-varying coefficients.

We perform a simulation study to cross-analyze survival data that follow either Aalen's linear model <sup>[3]</sup>, Cox's PH model <sup>[1]</sup> or Gray's TVC model <sup>[2]</sup> for the conditional hazard rate involving a single continuous covariate. Particular attention is given to data from a model with time-varying regression coefficients.

It is often the case that right-censored survival data are analyzed without prior verification of the assumptions of an assumed model. The effect of misspecifying the true model is assessed by estimating the power of Aalen's, Cox's and Gray's model, respectively, to detect an existing effect and by comparing the mean square error and bias of the conditional estimator of survival obtained for each of the three models used in analyzing the data. Formulas for the bias and the mean square error of the conditional survival distribution rely on the ability of the fitted model to estimate survival. In this part we therefore see a direct application of the concepts developed in the first part of the thesis. Specifically, the survival function estimator for Gray's PC-TVC model <sup>[27]</sup> (Valenta and Weissfeld, 2002) is now being used for the evaluation of the bias, mean square error and the true survival function when Gray's model is assumed.

Under the multiplicative Cox PH model, the effect of one subject's set of covariates  $z \in R^p$  on the conditional hazard rate  $\lambda(t|z)$  is assumed to be proportional to that of any other subject, and is defined as:

$$\lambda(t|z) = \lambda_0(t) \exp \{ \beta' z \}, \quad (24)$$

where  $\lambda_0(t)$  denotes the baseline hazard rate, which for the purpose of estimating the model parameters,  $\beta$ , is treated as a nuisance function. Parameter estimation for this model proceeds by maximizing the partial likelihood.

Gray's PC-TVC model is a generalization of the Cox PH hazards model which incorporates piecewise-constant time-varying regression coefficients. This allows for the modeling of right-censored survival data which violates the PH assumption. Another important distinction is that Gray's PC-TVC model also incorporates penalized splines in order to estimate the model parameters. Gray's PC-TVC model for the conditional hazard rate has the following form:

$$\lambda(t|z) = \lambda_0(t) \exp \{ \beta(t)' z \}, \quad (25)$$

where  $\beta(t) = \beta(\tau_j)$  for  $t \in [\tau_j; \tau_{j+1})$ ,  $j = 0, \dots, q$ . Here  $\tau_j$ ,  $j = 1, \dots, q$ , denote the internal knots,  $\tau_0 = 0$ , and  $\tau_{q+1} = T$  denotes the maximum observed (survival or censoring) time. Parameter estimates are obtained via maximization of the penalized partial likelihood (Gray <sup>[2]</sup>).

Aalen's additive model specifies a linear effect of the vector of subjects' covariates on

the conditional hazard rate:

$$\lambda(t|z) = \beta(t)'z, \quad (26)$$

where  $\beta(\cdot) \in R^{p+1}$  is a vector of time-varying coefficients and  $z \in R^{p+1}$  is a vector of covariate values with  $z_0 = 1$  being the intercept of the linear model. Note that in this formulation of Aalen's linear model  $\beta_0(t)$  represents a baseline hazard of the underlying population. The regression coefficients,  $\beta(t)$ , are usually estimated in their cumulative form under Aalen's model as

$$B_j(t) = \int_0^t \beta_j(y)dy, \quad j = 0, \dots, p. \quad (27)$$

The cumulative regression coefficients estimates  $\hat{B}(t) = (\hat{B}_0(t), \dots, \hat{B}_p(t))$  are obtained for each observed failure time except the last  $p + 1$  ordered cases. This is a convenient form allowing for a straightforward evaluation of the survival function, as shown in (28)

$$\begin{aligned} \hat{S}(t|z) &= \exp\left(-\int_0^t \lambda(y)dy\right) = \exp\left(-\int_0^t z' \hat{\beta}(y)dy\right) \\ &= \exp\left(z' \int_0^t \hat{\beta}(y)dy\right) = \exp\left(z' \hat{B}(t)\right). \end{aligned} \quad (28)$$

In this study we use the ordinary least squares estimates of  $\hat{B}(t)$  suggested by Aalen in 1989<sup>[3]</sup>. An apparent drawback of Aalen's estimating procedure is that it allows for negative estimates of the conditional hazard rate. Consequently, the estimated survival function based on Aalen's model is not necessarily a monotone function and therefore has to be corrected for non-monotonicity.

### 3.3 Bias and mean square error of the estimated survival

The bias of the survival function estimate at time  $t$  is defined as:

$$\text{Bias}(\hat{S}(t|z)) = \mathbb{E}(\hat{S}(t|z)) - S(t|z). \quad (29)$$

where  $\hat{S}(\cdot|\cdot)$  and  $S(\cdot|\cdot)$  denote the estimated and true survival function, respectively.

Consequently, the empirical bias is defined as:

$$\widehat{\text{Bias}}(\hat{S}(t|z)) = \frac{1}{n_s} \sum_{i=1}^{n_s} \hat{S}^{(i)}(t|z) - S(t|z), \quad (30)$$

where  $n_s$  represents number of simulations,  $\hat{S}^{(i)}(\cdot|\cdot)$  denotes the realization of survival function estimate from the  $i$ -th simulation,  $S(\cdot|\cdot)$  denotes the true survival function and  $z$  is a  $px1$  vector of covariate values at which the survival function is, conditionally, evaluated.

Using the notation in (29), the mean square error (MSE) of the estimated survival function evaluated at time  $t$  for a covariate vector  $z \in R^p$  is defined as follows:

$$\text{MSE}(\hat{S}(t|z)) = \mathbb{E}(\hat{S}(t|z) - S(t|z))^2. \quad (31)$$

Similarly, the empirical MSE calculated at time  $t$  for a covariate vector  $z \in R^p$  will have the form:

$$\widehat{\text{MSE}}(\hat{S}(t|z)) = \frac{1}{n_s} \sum_{i=1}^{n_s} (\hat{S}^{(i)}(t|z) - S(t|z))^2. \quad (32)$$

Results from the MSE analysis will be presented in relative form, i.e. as ratios of the MSE for the model under evaluation (say, Cox) and that from which the data were generated (e.g. Aalen).

### 3.4 Simulation study design

The simulation study we performed consists of four parts. In parts one (Section 3.8.1.) and two (Section 3.8.2) we study right-censored survival data generated from Aalen’s linear model for one continuous covariate. This case will be further referred to as the “Aalen Data Model”. In Sections 3.8.1. and 3.8.2. we examine separately the case of constant and time-varying regression coefficients for the Aalen Data Model. The time-varying coefficients in the Aalen Data Model are modeled using functions from the power family.

Part three (Section 3.8.3) features survival data generated from Gray’s TVC model (“Gray Data Model”). Here the time-varying coefficients were modeled using a general linear trend. Finally, in part four (Section 3.8.4) we assess data simulated from the Cox PH model (“Cox Data Model”) for one continuous covariate.

In each of the three components survival data were cross-analyzed assuming Aalen’s linear, Cox PH and Gray’s PC-TVC model, respectively. The power to detect the main effect, the bias and mean square error (MSE) of the survival estimate are reported for each of the three analysis models.

The bias and MSE were evaluated at 15 different time points and for 7 quantiles of the covariate distribution. The time points were selected as the following quantiles of the true survival distribution corresponding to the mean ( $z = 5$ ) of the covariate distrib-



ution: .99,.95,.90,.80,.75,.70,.60,.50,.40,.30,.25,.20,.10,.05,.01. Quantiles of the covariate distribution ( $N(5,1)$ ) for which the three survival function estimators were evaluated are the following: .01,.10,.25,.50,.75,.90,.99.

For Aalen's linear model we used 10 different statistics to test the significance of the main effect. They were obtained using different weight functions,  $K(\cdot)$ , which appear in the general formulae for the test statistic[39]:

$$U(t) = \int_0^t K(s)d\hat{B}(s) \quad (33)$$

Using the notation of Lee and Weissfeld[39], the weight functions,  $K(\cdot)$ , we used are the following: 1,  $K_2(t)$ ,  $K_2(t)\sqrt{Y_{\cdot}(t)}$ ,  $K_2(t)Y_{\cdot}(t)$ ,  $K_2(t)\sqrt{\hat{S}(t-)}$ ,  $K_2(t)\hat{S}(t-)$ ,  $\sqrt{Y_{\cdot}(t)}$ ,  $Y_{\cdot}(t)$ ,  $\sqrt{\hat{S}(t-)}$  and  $\hat{S}(t-)$ , where  $Y_{\cdot}(t)$  represents a number of patients being in the risk set at time  $t$  and  $\hat{S}(t)$  is a Kaplan-Meier survival estimate calculated at time  $t$ . These weight functions were designed specifically to detect different types of trends in the hazard[39]. We report the range of powers observed for the 10 statistics used in assessing the significance of the main effect when Aalen's analysis model is used.

For both Cox and Gray's model, the power to detect non-proportionality of the hazards is also reported. Results obtained from the simulation study are each based on 1000 simulations and a sample size of 300. They are presented for three different levels of censoring (light, moderate and heavy). The quality of the survival approximation obviously deteriorated with increasing level of censoring. Embedded in the corresponding formulas (to be introduced in the following paragraphs) is the asymptotic behavior of

simulated survival times resulting from increasing the size of the main effect. We are presenting the results for the size of the main effect which allows for enough variability in the generated survival times as well as for the manifestation of the effect of the time-varying regression coefficients where they are present.

For both Aalen's and Gray's model we have also considered other scenarios for modeling time-varying trends in the regression coefficients which lead to a closed form in evaluating both the simulation procedure and the formula for the true survival function given an assumed model. For the Aalen Data Model such suitable trends are obtained, for instance, using general linear, hyperbolic and piecewise constant functions or functions taking the form of some suitable probability density or a ratio of the probability density and cumulative density (e.g. using Cauchy distribution). For the Gray Data Model, other trends of choice would be e.g. a logarithmic or a piecewise-constant trend. The results we obtained for these other trends appeared to be consistent with those we obtained for our trends of choice.

### 3.5 Aalen Data Model

For the purpose of our simulations from Aalen's linear model with a single continuous covariate we assume that the baseline hazard is negligible (i.e. we set  $\beta_0(t) = 0$ ). We have generated right-censored survival data that follows Aalen's linear model with a single continuous covariate. The cases of constant and time-varying regression coefficients,  $\beta(t)$ , are considered separately. Time-varying coefficients are generated using functions from the power family:

$$\beta(t) = (t + c)^n, \quad (34)$$

where  $c > 0$ ,  $n \in N$ .

For  $z \in R^+$ , simulations of the survival data from the corresponding Aalen model proceed as follows:

$$-\log(1 - u) = \int_0^t z(y + c)^n dy = \frac{z}{n + 1} [(y + c)^{n+1}]_0^t = \frac{z}{n + 1} [(t + c)^{n+1} - c^{n+1}]_0^t, \quad (35)$$

where  $u$  is distributed according to a uniform  $U[0, 1]$  random variable.

Solving equation (35) for  $t$  renders the following solution:

$$t = \left( c^{n+1} - \frac{n + 1}{z} \log(1 - u) \right)^{\frac{1}{n+1}} - c. \quad (36)$$

The true survival function may in this case be evaluated as follows:

$$S(t|z) = \exp \left\{ -z \frac{(t + c)^{n+1} - c^{n+1}}{n + 1} \right\}. \quad (37)$$

In the case of a constant trend in the regression coefficient ( $\beta(t) = k, k > 0$ ), the formulae for generating survival times from the corresponding Aalen model takes on the following form:

$$t = -\log(1 - u)/(zk), \quad (38)$$

and the true survival function is evaluated as:

$$S(t|z) = \exp\{-zkt\}. \quad (39)$$

### 3.6 Gray Data Model

For Gray's TVC model the time-varying regression coefficients  $\beta(t)$  were modeled using a general linear trend  $\beta(t) = kt + c$ ,  $k \neq 0$ ,  $c \in R$ . Let  $u$  again represent a uniform  $U[0, 1]$  random variable. For  $z \neq 0$  and baseline hazard  $\lambda_0(t) = 1$ , the survival data generated from Gray's TVC model are simulated as follows:

$$-\log(1 - u) = \int_0^t e^{(ky+c)z} dy = \frac{1}{k} \int_c^{kt+c} e^{uz} du = \frac{e^{cz} (e^{tkz} - 1)}{kz}, \quad (40)$$

and, after a little algebra, we obtain

$$t = \frac{1}{kz} \log \left( 1 - e^{-cz} kz \log(1 - u) \right) . \quad (41)$$

The true survival function may then be evaluated as

$$S(t|z) = \exp \left\{ - \int_0^t e^{(ky+c)z} dy \right\} = \exp \left\{ - \left[ \frac{e^{(ky+c)z}}{kz} \right]_0^t \right\} = \exp \left\{ \frac{e^{cz} (1 - e^{tkz})}{kz} \right\}. \quad (42)$$

### 3.7 Cox Data Model

The main focus of this study is to assess ability of the different models to identify and correctly quantify the true nature of continuous effects with possibly time-varying coefficients when the true model for the data is either additive or multiplicative. Since under the Cox PH model the effects appearing in the model are assumed to remain constant over time, the Cox PH model covers the case of a multiplicative model where TVC are not present.

Assuming again that the baseline hazard  $\lambda_0(t)$  is set to 1, both the simulation procedure and the true survival function may be easily evaluated in the closed form.

### 3.8 Summary of the results

In each of the following four sections simulated right-censored survival data were cross-analyzed assuming either an additive or a multiplicative model for the hazard rate. Aalen's linear model was our choice for the class of additive models while the Cox PH model and Gray's PC-TVC model represented the class of multiplicative models. Both Aalen's linear model and Gray's PC-TVC model incorporate time-varying coefficients.

In all instances, Gray's PC-TVC model was fitted with 10 knots and 4 degrees of freedom for fitting piecewise-constant splines. Results we present are all based on 1000 simulations, a sample size of 300 and a variable level of censoring. They are shown for the size of the main effect allowing for both enough variability in simulated survival times (i.e. main effect significance) and a manifestation of the effect of time-varying coefficients,  $\beta(t)$ , where they are present.

The results are shown for a selected set of time points and a selected set of quantiles  $z$  of the covariate distribution. As described in Section 3.4, the time points were not chosen arbitrarily. Rather they are pre-selected probability quantiles of the true conditional survival distribution given the mean ( $z=5$ ) of the covariate distribution (see Section 3.4 for more details).

### 3.8.1 Aalen Data Model with TVC (power trend $\beta(t) = (t + 1)^3$ )

Survival data following the Aalen Data Model with one continuous covariate and time-varying coefficients of the form  $\beta(t) = (t + 1)^3$  were cross-analyzed assuming Aalen's linear, Cox PH and Gray's PC-TVC analysis model, respectively. Table 2 describes the true survival probabilities associated with selected time points and quantiles  $z$  of the covariate distribution  $N(5,1)$ .

The middle column of Table 2 presents the pre-selected set of survival probabilities given the mean covariate value. In other words, the set of time points listed in the first column of Table 2 are the probability quantiles of Section 3.4 for the mean of the covariate distribution ( $z=5$ ). As we move horizontally away from the middle column of Table 2 we observe how the true probability distribution varies marginally across different values of the covariate distribution  $N(5,1)$ . By applying the cubic trend to the range of survival times listed in the first column of Table 2 we may inspect the range of time-varying coefficients realized over the corresponding time points. We find that the regression coefficients ranged from 1 to about 3.2 which would seem to be sufficiently variable for our purposes, especially when this range is viewed relative to the specified size of the main effect.

Table 3 presents the power of the three analysis models to detect a single continuous effect when data are generated from the Aalen Data Model with TVC of the form  $\beta(t) = (t + 1)^3$ . Results are shown for a variable level of censoring. The power to detect



non-proportionality of the hazards when the assumed model is multiplicative is also reported based on the Cox and Gray analysis model. As mentioned earlier, the results we obtained for other suitable trends in  $\beta(t)$  appeared to be consistent with those we obtained for the power family. We considered three levels of censoring (light, moderate and heavy).

Results shown in Table 3 reveal that the Cox PH model had very good power to detect an existing main effect in Aalen's additive model for the hazard rate (Aalen Data Model). In fact, Aalen's model - the correct analysis model for the data at hand - had a lower power for all 10 statistics used when the censoring was light or moderate. We also observe that time-varying effects for the Aalen Data Model do not translate into non-proportional hazards when the data are being (incorrectly) analyzed as multiplicative. The power to detect non-proportional hazards is close to the size of the tests when the hazards are in fact proportional. This is obviously the reason why Gray's analysis model lagged behind the other two models in terms of power.

As shown in Figures 3 through 5, the bias of Aalen's analysis model increased linearly over time, while both multiplicative models rendered their maximum biases in the neighborhood of the median survival time. However, the absolute magnitude of the bias was otherwise comparable for all three analysis models. The relatively better performance of Aalen's analysis model in terms of bias however seemed to have deteriorated with increasing censoring. When comparing the performance of the individual models in

terms of both the bias and the MSE we focus on the times up to the censoring limit. It is because the survival times were not observable beyond that limit and the estimates of the bias and the MSE shown for these times are based on projecting the last available survival estimate into the range of times for which no (censored or un-censored) observations were available.

In terms of the MSE Aalen's analysis model generally outperformed both multiplicative models for the 25th, 50th and 99th percentile of the covariate distribution  $N(5,1)$ , yet only for times in the larger neighborhood of true median survival times. For the other percentiles of the covariate distribution and all corresponding time quantiles the MSE of both multiplicative models appeared to be smaller than that of Aalen's linear model - the correct analysis model. For the highest level of censoring (62 %), however, both multiplicative models did noticeably worse in terms of the MSE only for the 99th percentile of the covariate distribution.

**Table 2:** True survival probabilities for Aalen Data Model with TVC present (power trend  $\beta(t) = (t + 1)^3$ )

<b>Aalen Data Model, <math>\beta(t) = (t + 1)^3</math></b>							
Time points	<b>Quantiles of covariate distribution <math>z \sim N(5,1)</math></b>						
t	z(0.01)	z(0.10)	z(0.25)	z(0.50)	z(0.75)	z(0.90)	z(0.99)
0.002	0.995	0.993	0.991	0.990	0.989	0.987	0.985
0.010	0.973	0.963	0.957	0.950	0.943	0.938	0.928
0.020	0.945	0.925	0.913	0.900	0.887	0.876	0.857
0.042	0.888	0.847	0.824	0.800	0.776	0.756	0.721
0.053	0.857	0.807	0.780	0.750	0.721	0.697	0.656
0.065	0.826	0.767	0.735	0.700	0.667	0.639	0.593
0.089	0.761	0.684	0.643	0.600	0.560	0.526	0.473
0.117	0.690	0.597	0.549	0.500	0.455	0.419	0.362
0.147	0.613	0.506	0.453	0.400	0.353	0.316	0.261
0.184	0.525	0.408	0.353	0.300	0.255	0.220	0.171
0.205	0.476	0.357	0.301	0.250	0.207	0.175	0.131
0.230	0.423	0.302	0.248	0.200	0.161	0.132	0.095
0.298	0.292	0.180	0.136	0.100	0.073	0.055	0.034
0.358	0.202	0.108	0.075	0.050	0.033	0.023	0.012
0.471	0.085	0.033	0.019	0.010	0.005	0.003	0.001

**Table 3: Power Analysis for Aalen Data Model with TVC present**

(power trend  $\beta(t) = (t + 1)^3$ )

$\beta(t) = (t + 1)^3$ $z \sim N(5,1)$	Aalen Data Model					
	Aalen's Model		Cox Model		Gray's Model	
Censoring	Power (Range)	Power nPH	Power	Power nPH	Power	Power nPH
4.8%	.315 : .931	N/A	.934	.046	.871	.050
24.2%	.291 : .866	N/A	.868	.049	.790	.042
62.0%	.304 : .579	N/A	.577	.042	.450	.038

Aalen Data Model, 4.8% censoring,  $Z \sim N(5,1)$ ,  $\beta(y) = (y+1)^3$   
 Absolute Bias and MSE Relative to Aalen's Model

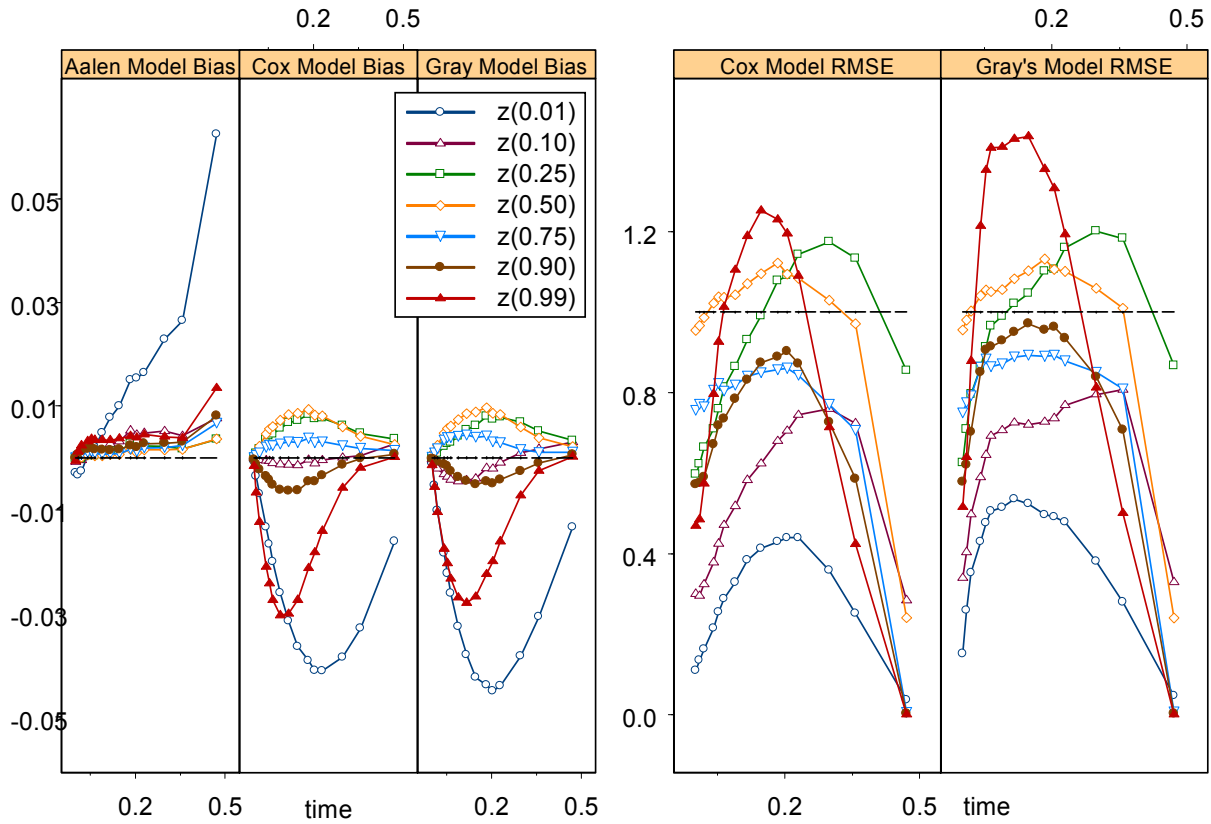


Figure 3: Bias and RMSE for Aalen Data Model with a Power Trend in  $\beta(y)$ , Light Censoring

Aalen Data Model, 24.2% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = (y+1)^3$   
 Absolute Bias and MSE Relative to Aalen's Model

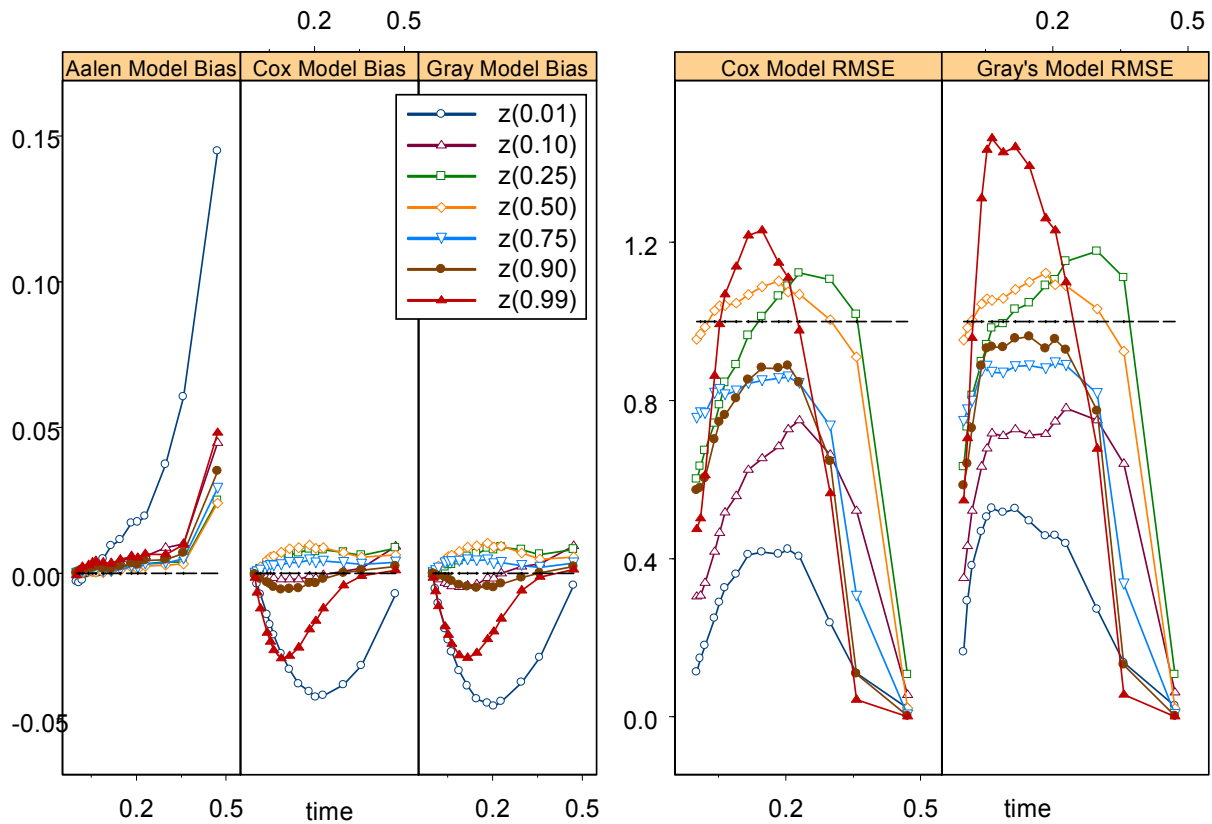


Figure 4: Bias and RMSE for Aalen Data Model with a Power Trend in  $\beta(y)$ , Moderate Censoring

Aalen Data Model, 62% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = (y+1)^3$   
 Absolute Bias and MSE Relative to Aalen's Model

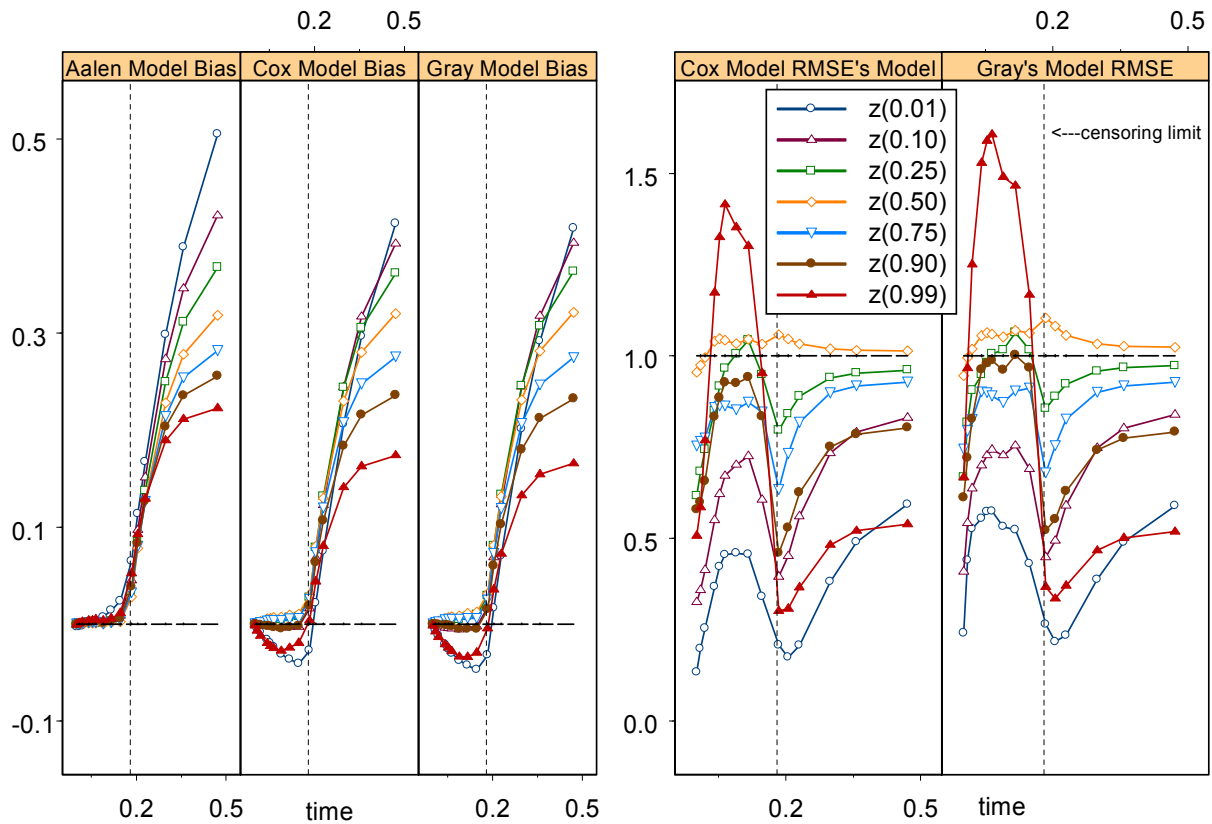


Figure 5: Bias and RMSE for Aalen Data Model with a Power Trend in  $\beta(y)$ , Heavy Censoring

### 3.8.2 Aalen Data Model with TVC absent (constant $\beta(t) = 1$ )

Here we follow a scenario analogous to that of the previous section. Tables 4 and 5 shown below relate to the additive Aalen Data Model in which the regression coefficient  $\beta(t)$  remains constant ( $\beta(t) = 1$ ). Table 4 describes the true survival probabilities for this model and Table 5 shows the power of the three analysis models to detect an existing main effect in the Aalen Data Model with a constant regression coefficient  $\beta(t)$ . The results are again shown for a variable level of censoring.

The power to detect non-proportionality of the hazards when the assumed model is multiplicative is also reported based on the Cox and Gray analysis models. The results we obtained in this section are very similar to those that we observed for the Aalen Data Model when time-varying coefficients were present. We may notice that for both multiplicative models the power to detect non-proportional hazards has slightly decreased in comparison with the previous results and consequently still remains in the margin for error. All of the patterns in the bias and MSE shown in Figures 6 through 8 remain very close to what we have observed in the previous section. While there is a small increase of scale in the bias for all three models, proportionally, the picture remains the same.

The true model for the data being analyzed again outperformed the two multiplicative models in terms of the MSE for the 25th, 50th and 99th percentile of the covariate distribution in the larger neighborhood of the median survival time. For all other quantiles of the covariate distribution and all time quantiles under the consideration both the Cox



PH and Gray's PC-TVC model performed better in terms of the MSE when compared to Aalen's linear model.

**Table 4:** True survival probabilities for Aalen Data Model with TVC absent (constant  $\beta(t) = 1$ )

Aalen Data Model, $\beta(t) = 1$							
Time points	Quantiles of covariate distribution $z \sim N(5,1)$						
t	z(0.01)	z(0.10)	z(0.25)	z(0.50)	z(0.75)	z(0.90)	z(0.99)
0.002	0.995	0.993	0.991	0.990	0.989	0.987	0.985
0.010	0.973	0.963	0.957	0.950	0.943	0.938	0.928
0.021	0.945	0.925	0.913	0.900	0.887	0.876	0.857
0.045	0.888	0.847	0.824	0.800	0.776	0.756	0.721
0.058	0.857	0.807	0.780	0.750	0.721	0.697	0.656
0.071	0.826	0.767	0.735	0.700	0.667	0.639	0.593
0.102	0.761	0.684	0.643	0.600	0.560	0.526	0.473
0.139	0.690	0.597	0.549	0.500	0.455	0.419	0.362
0.183	0.613	0.506	0.453	0.400	0.353	0.316	0.261
0.241	0.525	0.408	0.353	0.300	0.255	0.220	0.171
0.277	0.476	0.357	0.301	0.250	0.207	0.175	0.131
0.322	0.423	0.302	0.248	0.200	0.161	0.132	0.095
0.461	0.292	0.180	0.136	0.100	0.073	0.055	0.034
0.599	0.202	0.108	0.075	0.050	0.033	0.023	0.012
0.921	0.085	0.033	0.019	0.010	0.005	0.003	0.001

**Table 5: Power Analysis for Aalen Data Model with TVC absent ( $\beta(t) = 1$ )**

$\beta(t) = 1$ $z \sim N(5,1)$	Aalen Data Model					
	Aalen's Model		Cox Model		Gray's Model	
Censoring	Power (Range)	Power nPH	Power	Power nPH	Power	Power nPH
8.4%	.320 : .923	N/A	.923	.047	.851	.047
28.5%	.335 : .852	N/A	.849	.036	.758	.043
63.5%	.324 : .565	N/A	.561	.034	.428	.037

Aalen Data Model, 8.4% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 1$   
 Absolute Bias and MSE Relative to Aalen's Model

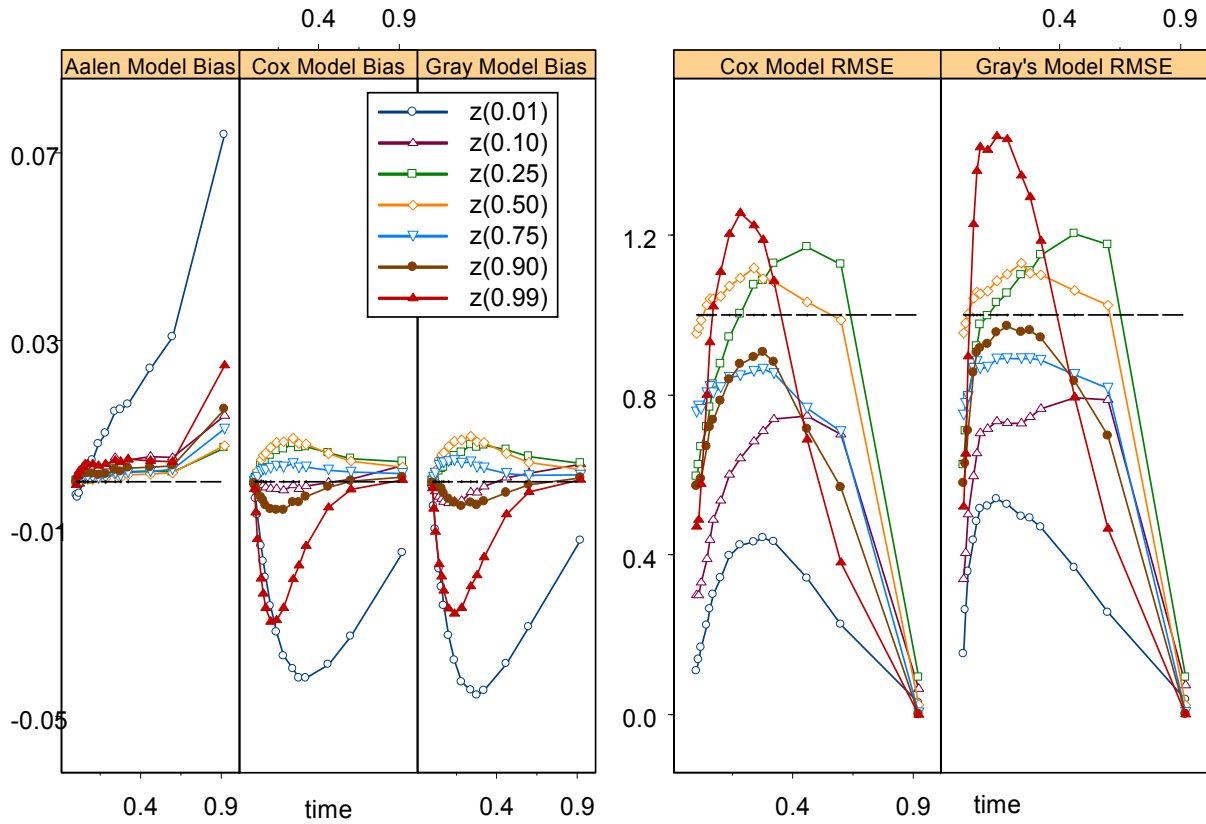


Figure 6: Bias and RMSE for Aalen Data Model with a Constant Trend in  $\beta(y)$ , Light Censoring

Aalen Data Model, 28.5% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 1$   
 Absolute Bias and MSE Relative to Aalen's Model

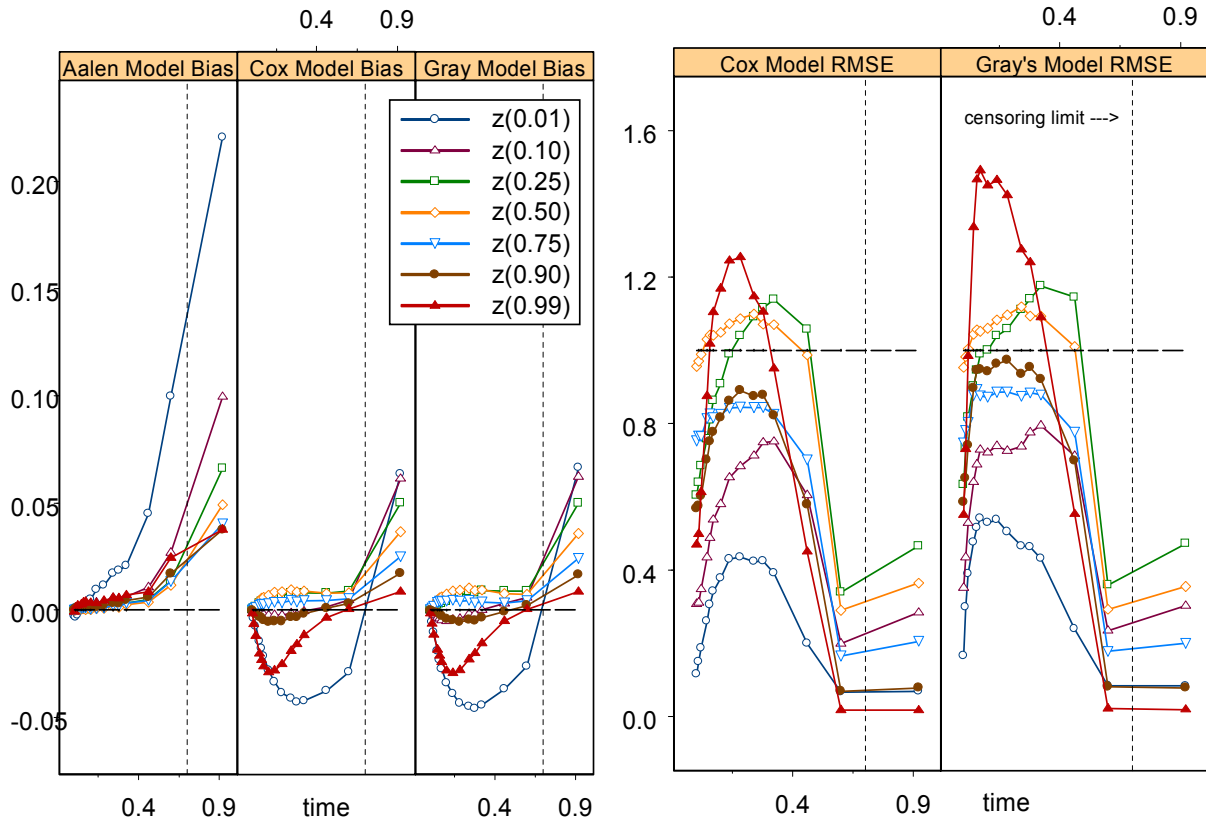


Figure 7: Bias and RMSE for Aalen Data Model with a Constant Trend in  $\beta(y)$ , Moderate Censoring

Aalen Data Model, 63.5% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 1$   
 Absolute Bias and MSE Relative to Aalen's Model

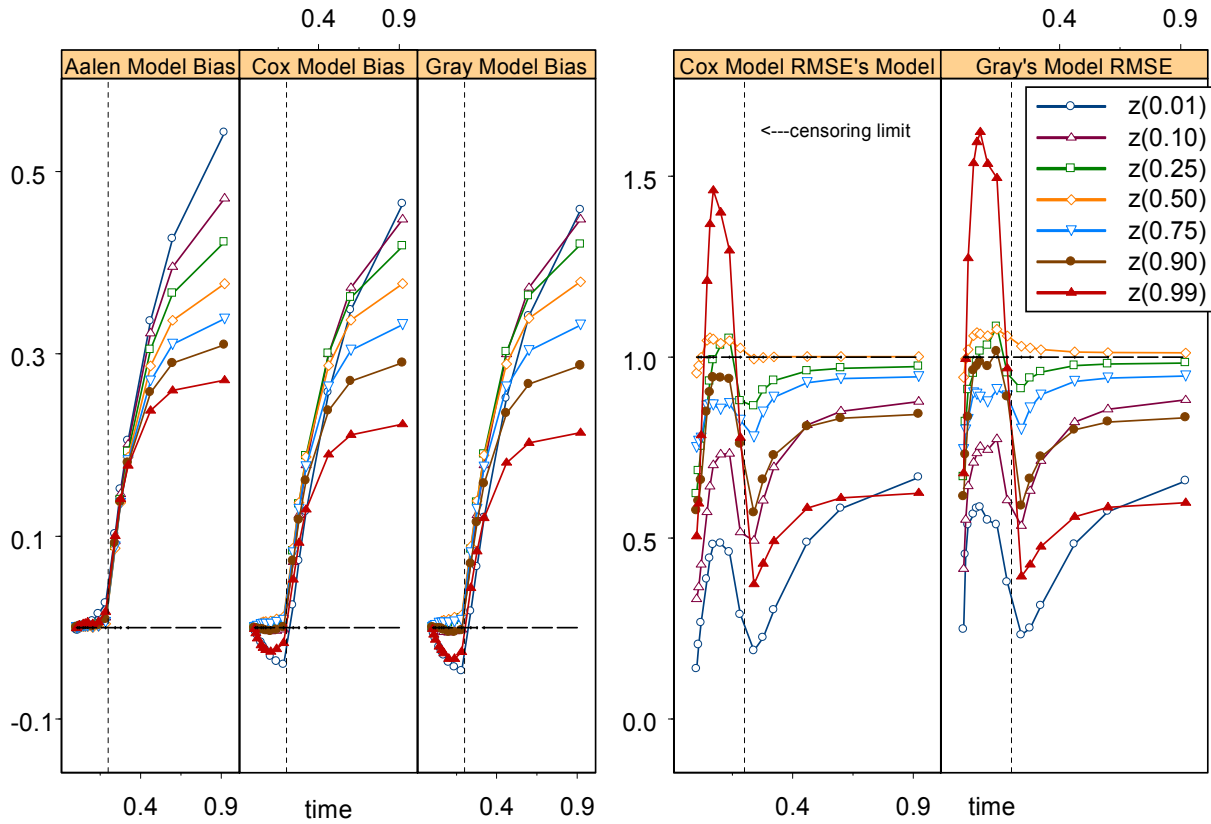


Figure 8: Bias and RMSE for Aalen Data Model with a Constant Trend in  $\beta(y)$ , Heavy Censoring

### 3.8.3 Gray Data Model with a linear trend in TVC ( $\beta(t) = 5t$ )

In this section we cross-analyze right-censored survival data generated from Gray's TVC model for a single continuous covariate assuming Aalen's linear, Cox PH and Gray's PC-TVC analysis models, respectively.

For the Gray Data Model we considered a linear trend in time-varying coefficients  $\beta(t) = 5t$  and the main effect generated from normal distribution with the mean 5 and standard deviation 1 (i.e.  $N(5,1)$ ). Note that there is an ambiguity in specifying the slope for the regression coefficient and the parameters of the main effect. In other words, Gray's TVC model specified by the regression coefficient  $\beta(t) = 5t$  and the main effect  $N(5,1)$  is identical to that for which  $\beta(t) = t$  and  $z \sim N(25,5^2)$ .

Table 6 describes the true survival probabilities corresponding to selected sets of time points and quantiles  $z$  of the covariate distribution  $N(5,1)$ . Table 7 presents the power of the three analysis models to detect an existing continuous effect in Gray Data Model where TVC are modeled using  $\beta(t) = 5t$ . The results are again shown for a variable level of censoring. The power to detect non-proportional hazards is reported based on the Cox and Gray's analysis models.

The power to detect a single continuous main effect from the Gray Data Model with time-varying coefficients being modeled using a linear trend  $\beta(t) = 5t$  was, as shown in Table 7, very high for all three analysis models under consideration. Both tests

of non-proportionality of the hazards had generally a good power when the censoring level was low or moderate and their power decreased with the introduction of heavier censoring. This was to be expected because the range of observed times to an event is inversely proportional to the level of censoring. As the range of the observed failure times decreases, the range of the corresponding values of time-varying coefficients and thus the power of the corresponding tests for non-proportionality of the hazard also decreases.

Gray's analysis model - the correct model for the data being analyzed - clearly outperformed the other two models in terms of bias and MSE as shown in Figures 9 through 11. Gray's analysis model maintained a very low level of the bias for all levels of censoring. The bias of the Gray's model remained very low in both absolute and relative terms. Here we should perhaps emphasize once more that we are primarily interested in how all the models perform (in terms of bias and MSE) for times not exceeding the censoring limit. For both Aalen's and Cox's analysis model the maximum bias was observed for times in the neighborhood of the median survival time corresponding to the mean of the covariate distribution.

The bias appeared to be smaller for Aalen's analysis model than for Cox analysis model, except for the case of two extreme covariates (1st and 99th percentile of the covariate distribution), when censoring was light or moderate. The reason is that Aalen's model incorporates time-varying coefficients and even though the model is not multiplicative,



its flexibility in terms of time-varying effects prevail over that of the Cox model when the main effects are not very extreme. Gray's TVC model formulae (25) and Aalen's linear model formulae (26) are actually related through the first order Taylor series approximation. Recalling the Gray's TVC model formulae (25) we can write:

$$\lambda(t|z) = \lambda_0(t) \exp \{ \beta(t)' z \} \approx \lambda_0(t) [1 + \beta(t)' z] = \lambda_0(t) + [\lambda_0(t) \beta(t)]' z \quad (43)$$

We observe that through the first order Taylor series approximation the baseline hazards of both models correspond directly while the regression coefficients from Aalen's linear model are obtained as a product of the baseline hazard evaluated at time  $t$  and time-varying regression coefficients from Gray's TVC model evaluated at time  $t$ .

This mathematical relationship between Aalen's linear and Gray's TVC model is the reason why under light and moderate censoring Aalen's analysis model seems to have performed better for Gray's model data in terms of the bias when compared to the Cox model, with the exception of extreme covariates (1% and 99% quantile of the covariate distribution). Since any linear model is concerned with predicting the mean, it is no surprise that the approximation failed for extreme covariates if we recall that the nature of the true underlying model for the data is not actually linear. With the introduction of the heavy censoring time-varying covariate effects effectively vanish and we no longer observe this phenomenon.

Inspecting the results for RMSE reveals that misspecifying Gray's TVC model - the true model for the data in this section - has far more serious consequences relative to what we have observed for the Aalen Data Model. Results related to the RMSE shown in Figures 9 through 11 are presented only for the range  $[0,50]$ . The actual range for RMSE associated with 99th percentile of the covariate distribution  $N(5,1)$  was  $[0,803]$  for the Cox analysis model and  $[0,44076]$  for Aalen's analysis model when censoring was light. Both ranges decreased with increasing level of censoring.

Inflation of the relative error associated with the 99th percentile of the covariate distribution underlines the great flexibility maintained by Gray's PC-TVC model over the whole range of covariate values. Aalen's analysis model seemed to have performed better than Cox model in terms of RMSE when censoring was light or moderate and the covariate values were not extreme. This again reflects a comparatively better adaptivity of Aalen's linear model to the presence of TVC than that of Cox PH model even though the class of the model is misspecified. Since the time-varying effects in our simulations diminished with increasing level of censoring, Cox analysis model again prevailed over Aalen's model in terms of RMSE when censoring increased to heavy.

**Table 6:** True survival probabilities for Gray Data Model with a linear trend in TVC ( $\beta(t) = 5t$ )

Gray Data Model, $\beta(t) = 5t$							
Time points	Quantiles of covariate distribution $z \sim N(5,1)$						
t	z(0.01)	z(0.10)	z(0.25)	z(0.50)	z(0.75)	z(0.90)	z(0.99)
0.009	0.991	0.990	0.990	0.990	0.990	0.990	0.989
0.033	0.959	0.955	0.953	0.950	0.947	0.944	0.938
0.052	0.928	0.917	0.909	0.900	0.889	0.879	0.858
0.075	0.878	0.848	0.827	0.800	0.768	0.735	0.668
0.084	0.856	0.816	0.787	0.750	0.706	0.660	0.567
0.092	0.835	0.785	0.748	0.700	0.643	0.585	0.468
0.105	0.795	0.723	0.670	0.600	0.519	0.437	0.288
0.116	0.756	0.661	0.591	0.500	0.398	0.302	0.148
0.127	0.716	0.597	0.510	0.400	0.284	0.185	0.059
0.137	0.674	0.528	0.424	0.300	0.181	0.095	0.015
0.143	0.650	0.490	0.378	0.250	0.135	0.060	0.006
0.149	0.624	0.449	0.330	0.200	0.094	0.034	0.002
0.163	0.557	0.348	0.219	0.100	0.029	0.005	0.000
0.173	0.505	0.275	0.148	0.050	0.009	0.001	0.000
0.190	0.416	0.167	0.062	0.010	0.000	0.000	0.000

**Table 7: Power Analysis for Gray Data Model with  $\beta(t) = 5t$**

$\beta(t) = 5t$ $z \sim N(5,1)$	Gray Data Model					
	Aalen's Model		Cox Model		Gray's Model	
Censoring	Power (Range)	Power nPH	Power	Power nPH	Power	Power nPH
7.6%	.793 : 1.00	N/A	1.00	.855	1.00	.805
22.8%	.769 : 1.00	N/A	1.00	.777	1.00	.715
62.4%	.687 : .982	N/A	.983	.384	.982	.377

It is important to note that extreme covariate values are often of interest in clinical practice. The expected probability of survival at some point in time given a patient's (extreme) covariate values may well take part in a clinical decision process. It may, for instance, serve as one of the indicators for assessing patient's eligibility for receiving a transplant. In other words, the ability to correctly quantify survival expectation for subjects with extreme covariates may play an important role in real life decision making.

This leads to a following - though perhaps almost trivial - summary of the findings from this section: there is no substitute to Gray's analysis model when it is the true model for the data. Substantial bias is observed in the neighborhood of median survival times for both alternative models and relative mean square errors tend to be inflated for extreme covariate values, which remains the case when Aalen's analysis model is used for all

levels of censoring. The magnitude of the observed error when the Gray Data Model was misspecified was, in terms of both bias and MSE, generally much higher than that observed in the case of the Aalen Data Model.

The powers recorded in Table 7 may perhaps give a wrong impression that the main effect from the Gray Data Model is almost equally likely to be identified using each of the three analysis models. It is not generally the case. If, for instance, the regression coefficients are such that their average over time is approximately zero than the power to identify the main effect using Cox analysis model will diminish significantly relative to using native Gray's analysis model. We illustrate these cases using the logarithmic and adjusted linear trend in  $\beta(t)$ , respectively. The corresponding powers are presented in Tables 8 and 9, respectively.

**Table 8:** Power Analysis for Gray Data Model with  $\beta(t) = \log(t)$

$\beta(t) = \log(t)$ $z \sim N(2.5,1)$	Gray Data Model					
	Aalen's Model		Cox Model		Gray's Model	
Censoring	Power (Range)	Power nPH	Power	Power nPH	Power	Power nPH
12.6%	.052 : .780	N/A	.320	1.00	1.00	1.00
25.4%	.062 : .691	N/A	.194	1.00	1.00	1.00
50.7%	.043 : .622	N/A	.078	.992	.981	.994

**Table 9: Power Analysis for Gray Data Model with  $\beta(t) = t - 1$**

$\beta(t) = t - 1$ $z \sim N(5,1)$	Gray Data Model					
	Aalen's Model		Cox Model		Gray's Model	
Censoring	Power (Range)	Power nPH	Power	Power nPH	Power	Power nPH
12.5%	.100 : .825	N/A	.754	.960	.990	.941
30.1%	.060 : .731	N/A	.582	.991	.948	.888
48.2%	.053 : .576	N/A	.351	.819	.836	.807

The results shown in Tables 8 and 9 document the built-in averaging effect of the Cox PH model expressed in the loss of power relative to Gray's PC-TVC model. They also support our previous finding that Aalen's linear model seems to be more adaptive to the presence of time-varying effects than the Cox PH model even when the true model for the data is multiplicative. Since the powers to detect the main effect from Gray Data Model shown in Tables 8 and 9 diminished for both Cox's and Aalen's analysis model relative to what we observe in Table 7, we may expect that the corresponding bias and RMSE will be of even greater magnitude for non-native models than that we report in Figures 9 through 11.

Gray Data Model, 7.6% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 5y$   
 Absolute Bias and MSE Relative to Gray's Model

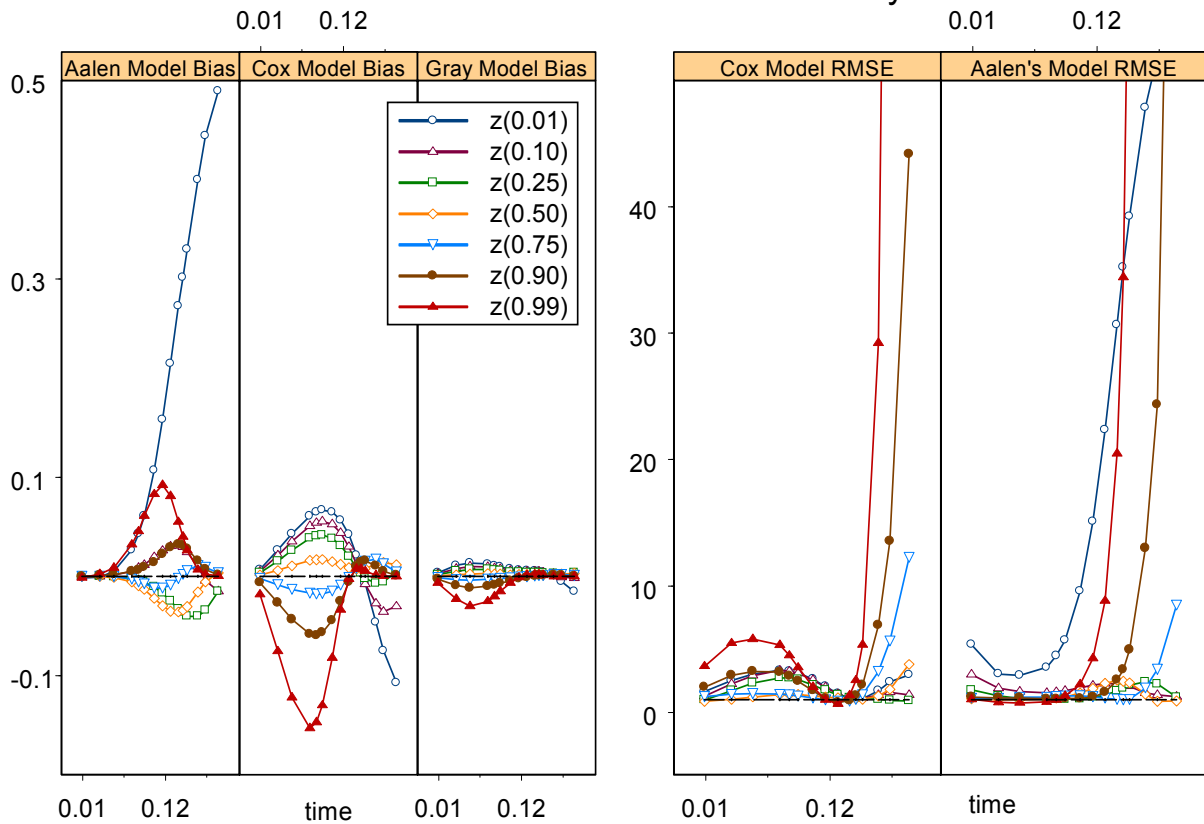


Figure 9: Bias and RMSE for Gray Data Model with a Linear Trend in  $\beta(y)$ , Light Censoring

Gray Data Model, 22.8% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 5y$   
 Absolute Bias and MSE Relative to Gray's Model

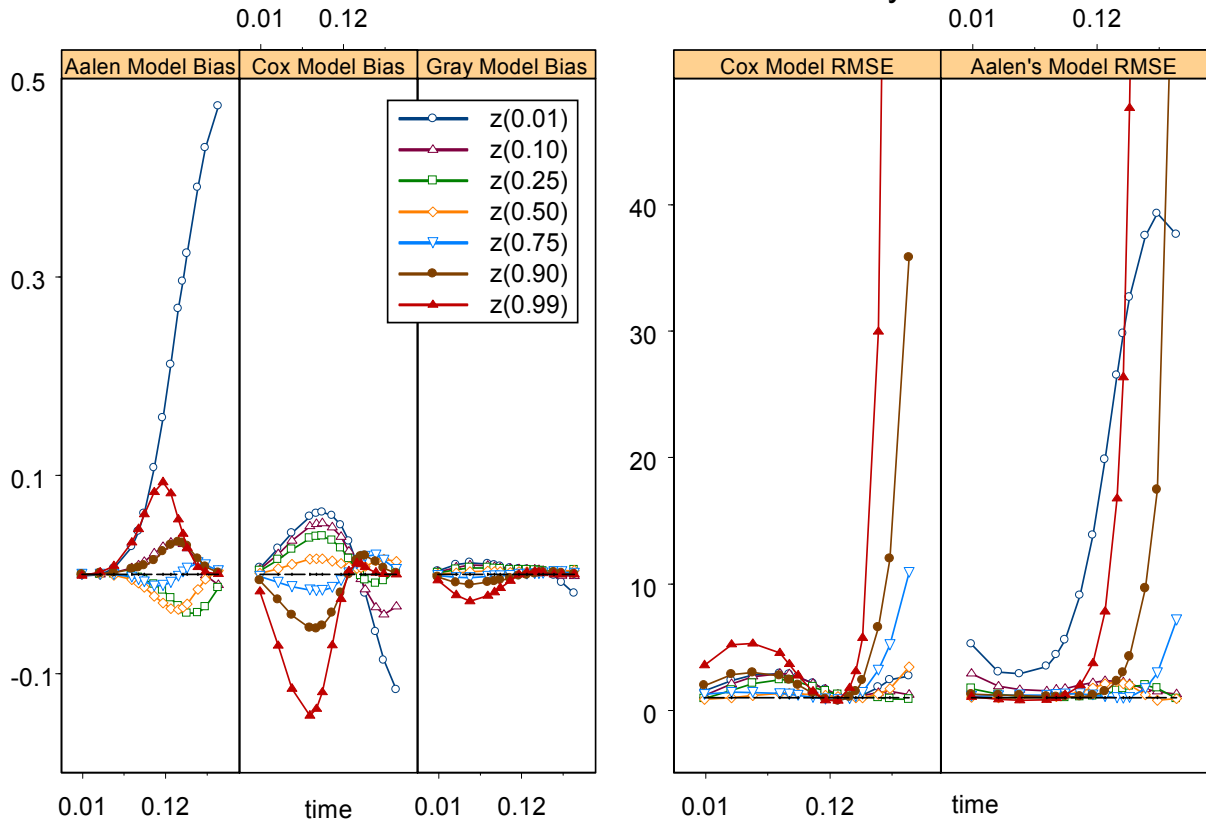


Figure 10: Bias and RMSE for Gray Data Model with a Linear Trend in  $\beta(y)$ , Moderate Censoring



Gray Data Model, 62.4% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 5y$   
 Absolute Bias and MSE Relative to Gray's Model

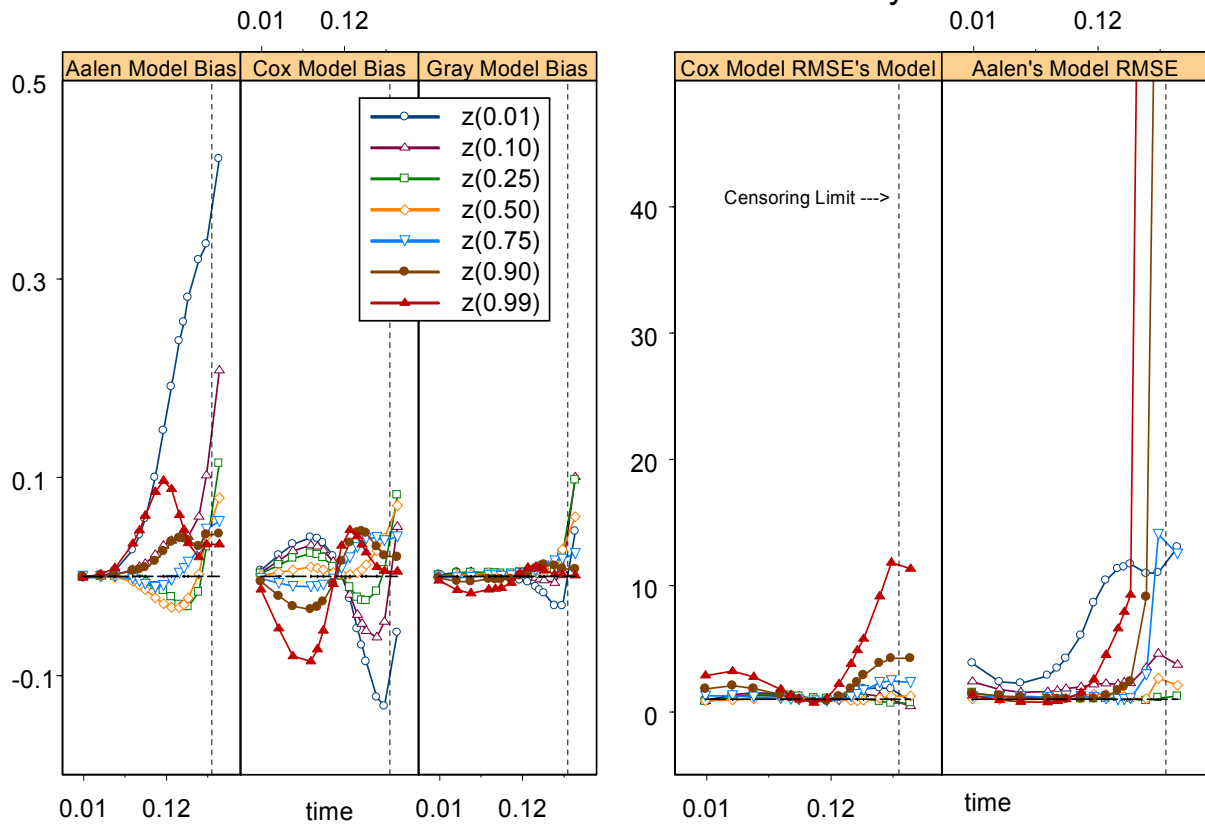


Figure 11: Bias and RMSE for Gray Data Model with a Linear Trend in  $\beta(y)$ , Heavy Censoring

### 3.8.4 Cox Data Model (constant $\beta(t) = 0.5$ )

Table 10 describes the true survival probabilities based on the Cox Data Model for a single continuous covariate with a constant regression coefficient  $\beta(t) = 0.5$ . They are shown for selected sets of time points and quantiles  $z$  of the covariate distribution  $N(5,1)$ .

Table 11 presents the power of the three analysis models to detect an existing continuous effect in the Cox Data Model with  $\beta(t) = 0.5$ . Results are again shown for a variable level of censoring. The size of the tests for non-proportionality of the hazards is also reported based on the Cox's and Gray's analysis models.

The results from Table 11 indicate that all three models maintained a very high power to detect an existing main effect from the Cox Data Model across all levels of censoring. The observed bias and MSE are smallest for the Cox analysis model - the true model for the data being analyzed. However, the error committed by the two alternative models differed dramatically. Gray's PC-TVC model seemed to render a reasonably small error in terms of both the bias and the MSE, especially when censoring was light or moderate. This should not be very surprising since Gray's PC-TVC model may be viewed as a piecewise PH model where the estimation is based on the penalized partial likelihood. On the other hand, the error observed for Aalen's analysis model appeared to be dramatic in terms of both the bias and MSE. The most serious bias was observed for Aalen's analysis model in the neighborhood of the median survival time corresponding to the mean of the covariate distribution.

The RMSE presented in Figures 12 through 14 is again restricted to the range  $[0,50]$ . The actual range of RMSE values from Aalen's analysis model generally exceeded the range of the values we have observed for Aalen's model in the previous section when TVC were present. For example, for a light level of censoring the actual range of RMSE values associated with the 99th percentile of the covariate distribution was  $[0, 2.8 * 10^6]$  for Aalen's analysis model, while for Gray's analysis model the range was  $[0,6.7]$ .

This seems to support our finding from the previous section that when time-varying effects are present, Aalen's analysis model may gain some flexibility even though the class of the model is not correct. The argument is actually made using a reversed implication: when TVC are not present in the true underlying multiplicative model for the data, Aalen's analysis model cannot materialize its flexibility towards time-varying effects and the observed MSE is of a greater magnitude than if TVC were present.

Finally, Table 11 reveals that the power of both tests detecting non-proportionality of the hazards remained appropriately within the margin for error. The size of the test for non-proportionality based on the Gray's routine appeared to be larger than that of the test implemented in Splus 2000 for Cox's PH model (Grambsch and Therneau <sup>[40]</sup>, 1984).

**Table 10:** True survival probabilities for Cox Data Model with  $\beta(t) = 0.5$

Cox Data Model, $\beta(t) = 0.5$							
Time points	Quantiles of covariate distribution $z \sim N(5,1)$						
t	z(0.01)	z(0.10)	z(0.25)	z(0.50)	z(0.75)	z(0.90)	z(0.99)
0.001	0.997	0.995	0.993	0.990	0.986	0.981	0.968
0.004	0.984	0.973	0.964	0.950	0.931	0.907	0.849
0.009	0.968	0.946	0.928	0.900	0.863	0.819	0.714
0.018	0.933	0.889	0.853	0.800	0.732	0.655	0.490
0.024	0.914	0.859	0.814	0.750	0.668	0.579	0.398
0.029	0.895	0.829	0.775	0.700	0.607	0.508	0.319
0.042	0.852	0.764	0.694	0.600	0.489	0.379	0.195
0.057	0.805	0.694	0.610	0.500	0.379	0.268	0.109
0.075	0.751	0.617	0.520	0.400	0.277	0.176	0.053
0.099	0.686	0.530	0.423	0.300	0.185	0.102	0.021
0.114	0.648	0.482	0.372	0.250	0.143	0.072	0.012
0.132	0.605	0.428	0.317	0.200	0.105	0.047	0.006
0.189	0.487	0.297	0.193	0.100	0.040	0.013	0.001
0.246	0.392	0.206	0.118	0.050	0.015	0.003	0.000
0.378	0.237	0.088	0.037	0.010	0.002	0.000	0.000

**Table 11:** Power Analysis for Cox Data Model with  $\beta(t) = 0.5$

$\beta(t) = 0.5$ $z \sim N(5,1)$	Cox Data Model					
	Aalen's Model		Cox Model		Gray's Model	
Censoring	Power (Range)	Power nPH	Power	Power nPH	Power	Power nPH
6.2%	.663 : 1.00	N/A	1.00	.040	1.00	.056
23.6%	.712 : 1.00	N/A	1.00	.035	1.00	.041
62.7%	.852 : .999	N/A	.999	.050	.996	.051

Cox Data Model, 6.2% censoring,  $Z \sim N(5,1)$ ,  $\beta(y) = 0.5$   
 Absolute Bias and MSE Relative to Aalen's Model

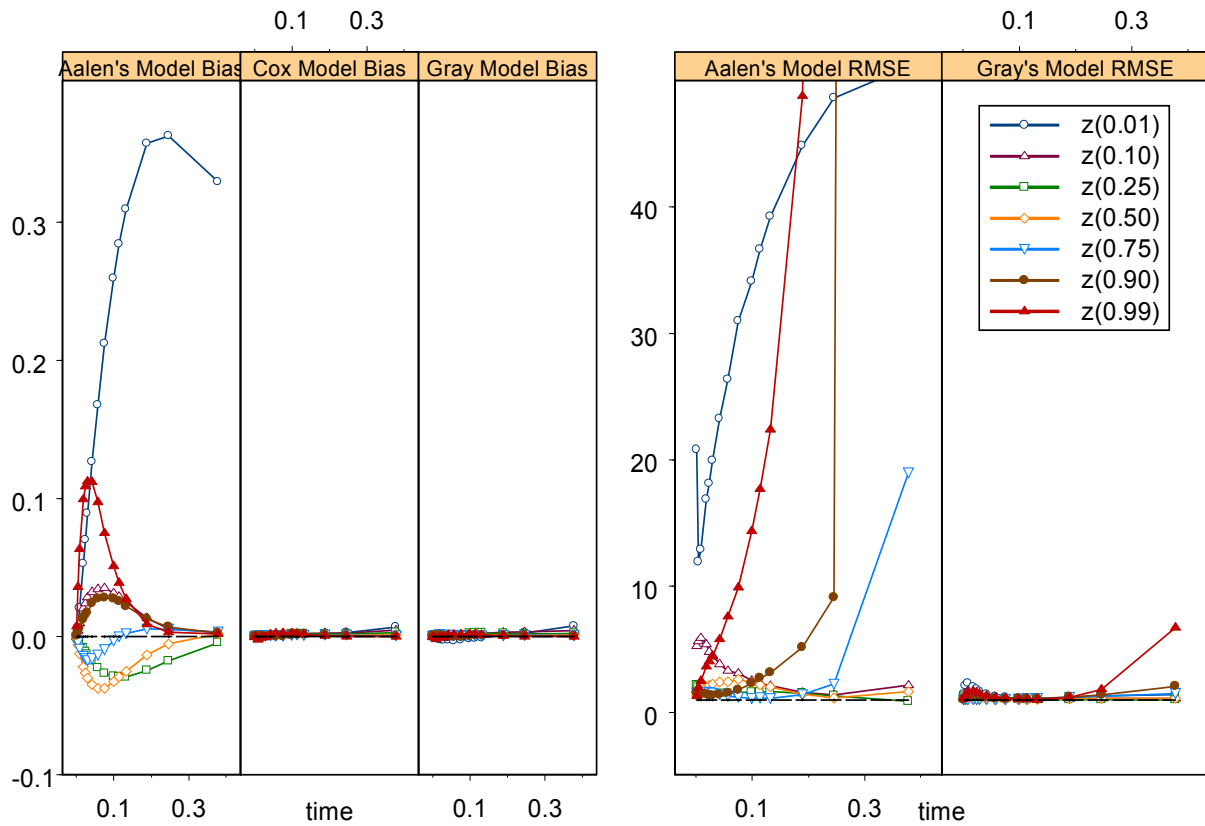


Figure 12: Bias and RMSE for Cox Data Model with a Constant Trend in  $\beta(y)$ , Light Censoring

Cox Data Model, 23.6% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 0.5$   
 Absolute Bias and MSE Relative to Aalen's Model

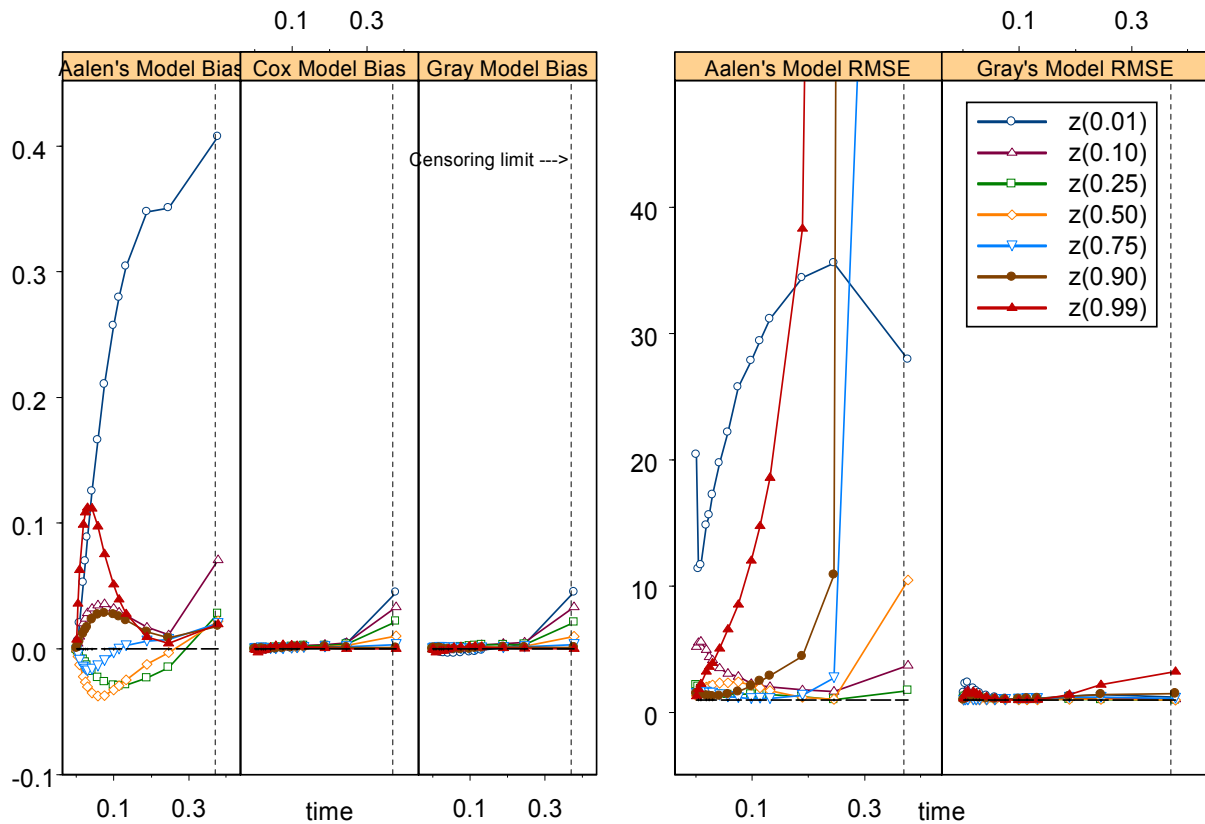


Figure 13: Bias and RMSE for Cox Data Model with a Constant Trend in  $\beta(y)$ , Moderate Censoring

Cox Data Model, 62.7% censoring,  $Z \sim N(5, 1)$ ,  $\beta(y) = 0.5$   
 Absolute Bias and MSE Relative to Aalen's Model

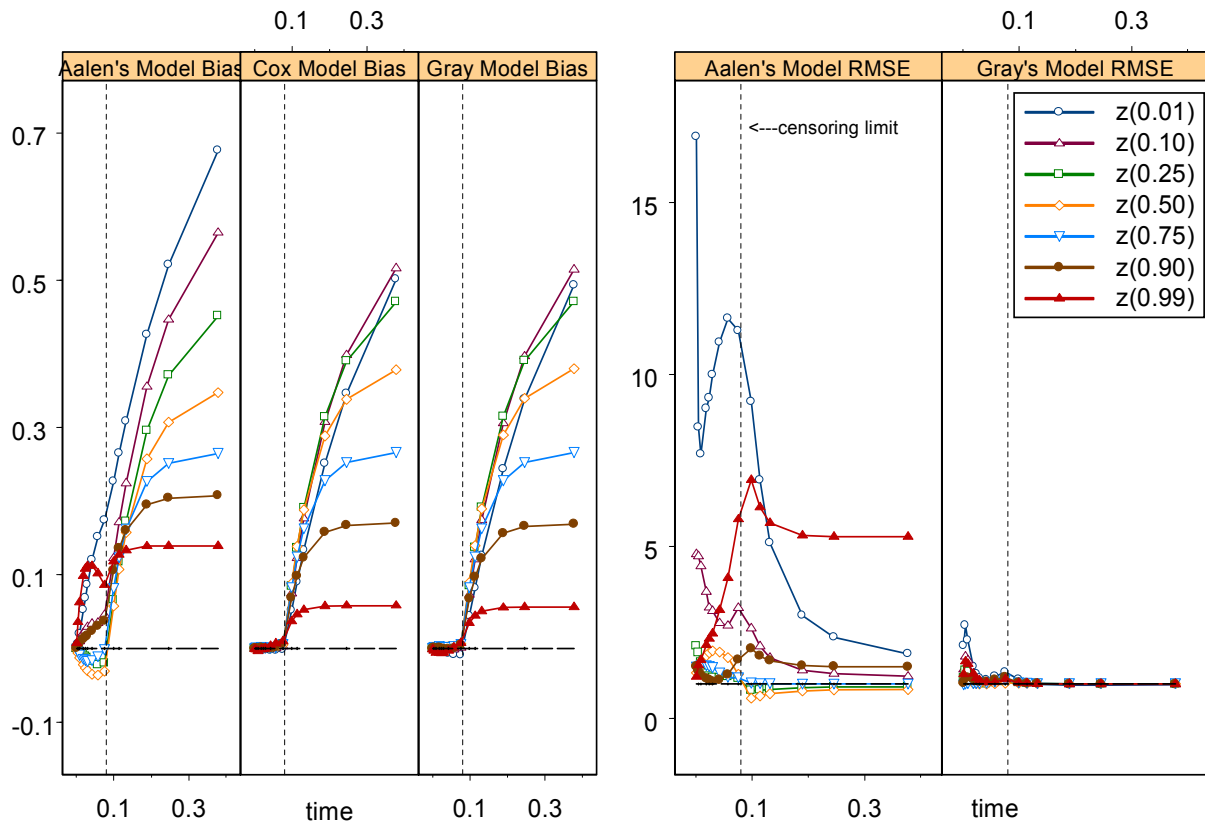


Figure 14: Bias and RMSE for Cox Data Model with a Constant Trend in  $\beta(y)$ , Heavy Censoring



### 3.9 Conclusions

The simulation study we performed indicates that when the true underlying model for right-censored survival data is multiplicative, misspecifying the class of the model at the analysis stage has far more serious consequences in terms of the observed bias and MSE of the estimates of the conditional survival distribution than when the scenario is reversed. In particular, when the survival data followed the Aalen Data Model we have observed that for the 1%, 10%, 75% and 90% quantiles of the covariate distribution the mean square errors of the survival estimates were smaller for both multiplicative models at all time points under consideration than those for the native Aalen's model. This was the case irrespective of whether the time-varying coefficients were present or not.

This was virtually never the case when the true underlying model for the data was multiplicative. In that case the correct analysis model for the data almost always outperformed the other non-native analysis models in terms of both bias and MSE. Consequences for misspecifying the multiplicative model for the data seemed to be generally far more serious as compared to what we have observed when the true underlying model for the data was additive. Aalen's model appeared to have gained some flexibility relative to the Cox analysis model when time-varying coefficients were present in the underlying multiplicative model for the data. Our results indicate, however, that there is no substitute for Gray's PC-TVC model when it is the native model for the data. The observed margin for error when the Gray Data Model is misspecified was the greatest of the four

cases we have studied in terms of both the bias and MSE. Results obtained from our simulation study support a case for virtual exclusivity of Gray's analysis model when the true underlying model for the data is multiplicative and time-varying coefficients are present.

### **Statement of Relevance**

As mentioned earlier, in practice it is often the case that right-censored survival data are being analyzed using generally available statistical software without prior verification of the assumptions of the underlying statistical model for the data. The simulation study we performed assessed the consequences when the assumptions are not met. The consequences were expressed in terms of the observed bias and mean square error of the conditional estimator of survival, and also in terms of power of the individual analysis models to detect an existing main effect in the model. We have observed that when the true model for the data is misspecified at the analysis stage, the magnitude of the observed error may be enormous (e.g. Gray Data Model). Conclusions of our study underline the importance of using appropriate rather than convenient statistical procedures when analyzing right-censored survival data.

### 3.10 Acknowledgments

We thank Dr. Robert J. Gray from Harvard University and Dana Farber Institute for making his own implementation of the modeling routine for Gray's time-varying coefficients model available to us. We also thank our colleague Zekarias T. Berhane for letting us use his own implementation of the OLS estimation procedure for Aalen's linear model. This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (grant NIDDK K25 DK59928).

## References

- [1] D. R. Cox. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Ser. B*, 34:187–220, 1972.
- [2] R. J. Gray. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association*, 87:942–951, 1992.
- [3] Odd O. Aalen. A linear regression model for the analysis of life times. *Statistics in Medicine*, 8:907–925, 1989.
- [4] P. K. Andersen, M. W. Bentzon, and J. P. Klein. Estimating the survival function in the proportional hazards regression model: A study of the small sample size properties. *Scandinavian Journal of Statistics*, 23:1–12, 1996.
- [5] J. D. Kalbfleisch and R. L. Prentice. Marginal likelihoods based on Cox’s regression and life model. *Biometrika*, 60:267–278, 1973.
- [6] B. Efron. The efficiency of Cox’s likelihood function for censored data. *Journal of the American Statistical Association*, 72:557–565, 1977.
- [7] K. E. Peace and R. E. Flora. Size and power assessments of tests of hypotheses on survival parameters. *Journal of the American Statistical Association*, 73:129–132, 1978.

- [8] J. H. Lubin. Analysis under Cox's failure time model using weighted least squares. *Biometrics*, 36:307–312, 1980.
- [9] D. Schoenfeld. Chi-squared goodness-of-fit tests for the proportional hazards regression model. *Biometrika*, 67(1):145–153, 1980.
- [10] D. Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69(1):239–241, 1982.
- [11] A. A. Tsiatis. A large sample study of Cox's regression model. *The Annals of Statistics*, 9(1):93–108, 1981.
- [12] P. K. Andersen and R. D. Gill. Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, 10:1100–1120, 1982.
- [13] M. E. Johnson, H. D. Tolley, M. C. Bryson, and A. S. Goldman. Covariate analysis of survival data: A small-sample study of Cox's model. *Biometrics*, 38:685–698, 1982.
- [14] W. H. Stewart and D. A. Pierce. Efficiency of Cox's model in estimating regression parameters with grouped survival data. *Biometrika*, 69:539–545, 1982.
- [15] K. R. Bailey. The asymptotic joint distribution of regression and survival parameter estimates in the Cox regression model. *The Annals of Statistics*, 11:39–48, 1983.

- [16] K. R. Bailey. Asymptotic equivalence between the Cox estimator and the general ML estimators of regression and survival parameters in the Cox model. *The Annals of Statistics*, 12:730–736, 1984.
- [17] S. M. Gore, S. J. Pocock, and G. R. Kerr. Regression models and non-proportional hazards in the analysis of breast cancer survival. *Applied Statistics*, 33(2):176–195, 1984.
- [18] C. L. Link. Confidence intervals for the survival function using Cox’s proportional-hazard model with covariates. *Biometrics*, 40:601–609, 1984.
- [19] R. D. Gill. Understanding Cox’s regression model: A martingale approach. *Journal of the American Statistical Association*, 79:441–447, 1984.
- [20] D. Gamerman and M. West. An application of dynamic survival models in unemployment studies. *The Statistician*, 36:269–274, 1987.
- [21] 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.
- [22] F. O’Sullivan. Nonparametric estimation in the Cox model. *The Annals of Statistics*, 21(1):124–145, 1993.
- [23] R. J. Gray. Spline-based tests in survival analysis. *Biometrics*, 50:640–652, 1994.

- [24] T. J. Hastie and R. J. Tibshirani. Varying coefficient models (with discussion). *Journal of the Royal Statistical Society, Ser. B*, 55(4):757–796, 1993.
- [25] 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.
- [26] A. N. Pettitt and I. Bin Daud. Investigating time dependence in Cox’s proportional hazards model. *Applied Statistics*, 39:313–329, 1990.
- [27] Zdenek Valenta and Lisa A. Weissfeld. Estimation of the survival function for the Gray’s piecewise-constant time-varying coefficients model. *Statistics in Medicine*, 21:717–727, 2002.
- [28] D. D. Cox and F. O’Sullivan. Asymptotic analysis of penalized likelihood and related estimators. *The Annals of Statistics*, 18(4):1676–1695, 1990.
- [29] Hess KR. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in Medicine*, 13:1045–1062, 1994.
- [30] P. J. M. Verweij and H. C. van Houwelingen. Time-dependent effects of fixed covariates in Cox regression. *Biometrics*, 51(1):1550–1556, 1995.
- [31] C. De Boor. *A Practical Guide to Splines*. Springer Verlag, New York, 1978.

- [32] T. R. Fleming and D. P. Harrington. *Counting Processes and Survival Analysis*, volume 1, section 4.3, pages 152, formula(3.29). John Wiley and Sons, Inc., One Wiley Drive, Somerset, NJ 08875, U.S.A., second edition, 1991.
- [33] D. R. Cox and N. Reid. Parameter orthogonality and approximate conditional inference (C/R: p18-39). *Journal of the Royal Statistical Society, Series B, Methodological*, 49:1–18, 1987.
- [34] Patricia J. Solomon. Effect of misspecification of regression models in the analysis of survival data (Corr: V73 p245). *Biometrika*, 71:291–298, 1984.
- [35] J. L. Hutton and P. J. Solomon. Parameter orthogonality in mixed regression models for survival data. *Journal of the Royal Statistical Society, Series B, Methodological*, 59:125–136, 1997.
- [36] C. A. Struthers and J. D. Kalbfleisch. Misspecified proportional hazard models. *Biometrika*, 73:363–369, 1986.
- [37] Nils Lid Hjort. On inference in parametric survival data models. *International Statistical Review*, 60:355–387, 1992.
- [38] Yi-Hwei Li, John P. Klein, and M. L. Moeschberger. Effects of model misspecification in estimating covariate effects in survival analysis for small sample sizes. *Computational Statistics and Data Analysis*, 22:177–192, 1996.



- [39] Eunyoung Lee and Lisa A. Weissfeld. Assessment of covariate effects in Aalen's additive hazard model. *Statistics in Medicine*, 17:983–998, 1998.
- [40] Patricia M. Grambsch and Terry M. Therneau. Proportional hazards tests and diagnostics based on weighted residuals (Corr: 95V82 p668). *Biometrika*, 81:515–526, 1994.
- [41] T. J. Hastie and R. J. Tibshirani. *Generalized Additive Models*. Chapman and Hall/CRC, 2000 N.W. Corporate Blvd., Boca Raton, Florida 33431, U.S.A., first CRC edition, 1999.
- [42] N. E. Breslow. Covariance analysis of censored survival data. *Biometrics*, 30:89–99, 1974.
- [43] S. R. Lipsitz and M. Parzen. A jackknife estimator of variance for Cox regression for correlated survival data. *Biometrics*, 52:291–298, 1996.
- [44] L. Marzec and Marzec P. On fitting Cox's regression model with time-dependent coefficients. *Biometrika*, 84(4):901–908, 1997.
- [45] L. J. Wei, D. Y. Lin, and L. Weissfeld. Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *Journal of the American Statistical Association*, 84(408):1065–1073, 1989.

- [46] F. O’Sullivan. Nonparametric estimation of relative risk using splines and cross-validation. *SIAM Journal on Scientific and Statistical Computing*, 9(3):531–542, 1988.
- [47] 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.
- [48] D. Y. Lin and T. R. Fleming. Confidence bands for survival curves under the proportional hazards model. *Biometrika*, 81(1):73–81, 1994.
- [49] R. J. Gray. Tests for variation over groups in survival data. *Biometrics*, 90(429):199–203, 1995.
- [50] D. Burr. A comparison of certain bootstrap confidence intervals in the Cox model. *Journal of the American Statistical Association*, 89(428):1290–1302, 1994.
- [51] A. Buja, T. Hastie, and R. Tibshirani. Linear smoothers and additive models. *The Annals of Statistics*, 17(2):453–555, 1989.
- [52] P. Sasieni. Maximum weighted partial likelihood estimators for the Cox model. *Journal of the American Statistical Association*, 88:144–152, 1993.
- [53] M. Schemper. Cox analysis of survival data with non-proportional hazard functions. *The Statistician*, 41:455–465, 1992.

- [54] L. J. Wei. The accelerated failure time model: A useful alternative to the Cox regression model in survival analysis (disc: P1881-1885). *Statistics in Medicine*, 11:1871–1879, 1992.
- [55] T. M. Therneau, P. M. Grambsch, and T. R. Fleming. Martingale-based residuals for survival models. *Biometrika*, 77:147–160, 1990.
- [56] R. J. Gray. Some diagnostic methods for Cox regression models through hazard smoothing. *Biometrics*, 46:93–102, 1990.
- [57] I. Annesi, T. Moreau, and J. Lellouch. Efficiency of the logistic regression and Cox proportional hazards models in longitudinal studies. *Statistics in Medicine*, 8:1515–1521, 1989.
- [58] D. Y. Lin and L. J. Wei. The robust inference for the Cox proportional hazards model. *Journal of the American Statistical Association*, 84:1074–1078, 1989.
- [59] J. Bretagnolle and C. Huber-Carol. Effects of omitting covariates in Cox’s model for survival data. *Scandinavian Journal of Statistics*, 15:125–138, 1988.
- [60] E. Arjas and D. Venzon. A test for discriminating between additive and multiplicative relative risks in survival analysis. *Applied Statistics*, 37:1–11, 1988.
- [61] J. Mau. On a graphical method for the detection of time-dependent effects of covariates in survival data. *Applied Statistics*, 35:245–255, 1986.

- [62] Fred W. Huffer and Ian W. McKeague. Weighted least squares estimation for Aalen's additive risk model. *Journal of the American Statistical Association*, 86:114–129, 1991.