# A comparison of statistical models for clustered survival data in multicenter clinical trials

Yaeji Lee

The Graduate School
Yonsei University
Department of Biostatistics and Computing

# A comparison of statistical models for clustered survival data in multicenter clinical trials

A Master's Thesis
Submitted to the Department of Biostatistics and Computing
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Science

Yaeji Lee

June 2021

This certifies that the master's thesis of ***Yaeji Lee*** is approved.

_____

***Chung Mo Nam***: Thesis Supervisor

_____

***Inkyung Jung***: Thesis Committee Member #1

_____

***Youn Nam Kim***: Thesis Committee Member #2

The Graduate School

Yonsei University

June 2021

# Table of Contents

# List of Tables

# List of Figures

# Abstract

Conducting multinational or multicenter clinical trials is a popular method of testing drug efficacy. In many of these studies, there are a large number of centers, but only a few patients are from each center and the ratio of the treatment group size to the control group size in a given stratum is imbalanced. Thus, such data should be analyzed with an appropriate model. Studies that include multiple centers should include stratified analyses. When the functional forms of the baseline hazards are unknown for each center, Glidden (2004) showed that the marginal hazard ratio is biased in a Cox proportional hazards model. There may be differences between subjects, such as those treated at different centers of a multicenter trial. In some instances, accounting for these differences is an integral part of conducting an appropriate analysis (ICH-E9). Thus, more complex methods are required to analyze these patient subgroups.

In this situation, a stratified cox model or frailty model can be considered. A stratified cox model stratifies by center, considering each center as a stratum. A frailty model expresses the effects of each center as a random effect, since patients in the same center are more similar to each other than to those in other centers. The aforementioned two models were compared under various scenarios. In a multicenter setting, the scenarios considered were: 1) the total sample size and the number of stratums being different, 2) the ratio of treatment to control in each center being different, 3) the functional form of the baseline hazards varying from center to center, and 4) the censoring percentage being different.

Under each scenario, the statistical performance of the two models were compared.

This study had two main objectives. The first objective was to check why unstratified analyses produce highly biased estimates in situations when stratified analyses are the correct procedures. The second objective was to determine whether the stratified Cox model or the random effects model is more appropriate for conducting multicenter analyses.

In conclusion, we attempt to present which model would be more appropriate in which context by comparing statistical models through simulations. This study is expected to provide an overall understanding of models that can be considered when conducting multinational or multicenter clinical trials, as well as help model selection.

---

Keywords: survival analysis, Cox proportional hazards model, stratified Cox model, gamma frailty model, multicenter clinical trial

# Chapter 1

# Introduction

Multicenter studies are clinical trials conducted simultaneously in various centers under one common clinical trial plan. Such studies are often conducted when it is too difficult for a single center to recruit enough participants. These allow sufficient subjects to be recruited to achieve the objective of the trial within a reasonable time-frame. Furthermore, it can provide a better basis for generalization of subsequent research results. For this reason, multicenter studies have become increasingly popular.

Most commonly, survival data processed through a proportional hazards model widely known as Cox regression model is first introduced by Cox (1972). The purpose of this model is to simultaneously evaluate the effects of multiple factors on survival. However, when the functional form of baseline hazards is unknown for each center, the marginal hazard ratio is biased in a Cox proportional hazards model. Colosimo (2001) showed that unstratified analyses produce significantly biased results when a stratified analysis should

have been used. Struthers and Kalbfleisch (1986) showed for a general situation of model misspecification that estimates for the proportional hazards model, converges in probability to a well-defined constant vector. Thus, more complex methods are required to analyze such clustered data of patients. Generally, individuals within a center are more similar than those in different centers. Thus, observations within a center may be correlated. If these similarities are sufficiently powerful, inferences ignoring the clustering can be seriously misleading.

The two main statistical models used in multicenter studies for considering the effects of centers are the stratified Cox proportional hazards model and the frailty model. The stratified Cox proportional hazards models are a useful extension of the Cox proportional hazards models to allow for adjusting covariates. It is used to adjust the effects of centers in multicenter trials, and the stratified Cox model treats centers as strata, modeling baseline hazards that differ from center to center. In this model, the regression coefficients are assumed to be the same in each stratum although the baseline hazard functions may be different and completely unrelated.

The frailty model expresses the effects of each center as random effects. Individuals may differ for survival probabilities in relation to hazards. Patients in some centers may have a greater risk of death than patients in others, and are therefore more likely to die faster. They may be described as being more frail, and variation in survival times can be explained by variability in this frailty effect across centers. These effects can be modeled by assuming that the subjects within a group share the same frailty.

When the stratified Cox model or frailty model is a true model, we would like to summarize why biases occur if it is not considered and to provide guidelines for which models would be appropriate for which situations through simulation of two models in various situations.

The rest of this paper is organized as follows. Section 1 explains the purpose of this study. Section 2 provides definitions for the models that were analyzed in this study. Section 3 provides a mathematical explanation for why unstratified analyses produce biased results when a stratified analysis should have been used. Section 4 evaluates the models' performance in various situations in terms of their bias and standard error through simulations. Section 5 discusses using the methods to analyze data. Section 5 presents conclusions and implications of this study.

# Chapter 2

# Methods and Materials

## 2.1 Cox's Proportional Hazards Models

Cox's proportional hazards model expresses the relationship between survival time $(T)$ and a set of explanatory variables through a hazard function. The hazard function $h(t)$, instantaneous rate of occurrence of the event, is expressed as a probability density function $f(t)$ at a given time $t$ divided by the survival function $S(t)$:

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

The survival function $S(t)$, which gives the probability that the event of interest has

not occurred by any specified time $t$, can be expressed using the cumulative hazard function:

$$H(t) = \int_0^t h(u)\,du$$

$$S(t) = \exp\left[-\int_0^t h(u)\,du\right] = \exp[-H(t)]$$

Thus, Cox's proportional hazards model for the $j$th individual at time $t$ is as follows:

$$h(t|Z_j) = h_0(t)\exp(\beta' Z_j)$$

In this model, $h_0(t)$ is the baseline hazard function, $\beta = (\beta_1, \ldots, \beta_p)'$ is a parameter vector. It is regression coefficients that estimate the effect of the covariate $Z$ of the $j$th individual. $Z_j = (Z_{j1}, \ldots, Z_{jp})'$ is the vector of covariates for the $j$th individual.

The Cox model performs inference through partial likelihood, the likelihood is calculated when individual $j$ fails. In the data, $j$th individual was characterized by $(T_j, \delta_j, Z_j(t))$, $j = 1, \ldots, n$ where $T_j$ was the time to event for the $j$th individual,

$\delta_j$ was the event indicator for the $j$th individual, where $\delta_j = 1$ if the event was observed and $\delta_j = 0$ if the response was censored; and $Z_j(t) = \left(Z_{j1}(t), \ldots, Z_{jp}(t)\right)'$ was the vector of covariates for the $j$th individual at time $t$. Given $Z_j$, censoring was non-informative, so the event and censoring time for the $j$th individual were assumed to be independent. In the case that there are no ties between the event times, $t_1 < t_2 < \cdots < t_D$ denoted the ordered observed event times and $Z_{(\ell)k}$ was the $k$th covariate associated with the observation whose failure time was $t_\ell$. The risk set at time $t_\ell$, $R(t_\ell)$, was the set of all observations which are survived without incident until just before time $t_\ell$. Thus, the partial likelihood function was expressed as:

$$L(\beta) = \prod_{\ell=1}^{D} \frac{\exp\left[\sum_{k=1}^{p} \beta_k Z_{(\ell)k}\right]}{\sum_{j \in R(t_\ell)} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]}$$

$LL(\beta) = \log[L(\beta)]$, so $LL(\beta)$ could be written as:

$$LL(\beta) = \sum_{\ell=1}^{D} \sum_{k=1}^{p} \beta_k Z_{(\ell)k} - \sum_{\ell=1}^{D} \log\left[\sum_{j \in R(t_\ell)} \exp\left(\sum_{k=1}^{p} \beta_k Z_{jk}\right)\right]$$

The regression coefficient $\beta$ was estimated from the score equations produced by the logarithmic partial likelihood function. Thus, the score equations of the Cox regression were written as follows:

$$U_h(\beta) = \frac{\partial LL(\beta)}{\partial \beta_h} = \sum_{\ell=1}^{D} Z_{(\ell)h} - \sum_{\ell=1}^{D} \frac{\sum_{j \in R(t_\ell)} Z_{jh} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]}{\sum_{j \in R(t_\ell)} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]} \quad (h = 1, \dots, p)$$

The information matrix is the negative of the matrix of second derivatives of the log-likelihood and is given by $I(\beta) = \left[I_{gh}(\beta)\right]_{p \times p}$ with $(g, h)$th element expressed as:

$$I_{gh}(\beta) = \sum_{\ell=1}^{D} \frac{\sum_{j \in R(t_\ell)} Z_{jg} Z_{jh} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]}{\sum_{j \in R(t_\ell)} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]} - \sum_{\ell=1}^{D} \frac{\sum_{j \in R(t_\ell)} Z_{jg} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]}{\sum_{j \in R(t_\ell)} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]} \frac{\sum_{j \in R(t_\ell)} Z_{jh} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]}{\sum_{j \in R(t_\ell)} \exp\left[\sum_{k=1}^{p} \beta_k Z_{jk}\right]}$$

## 2.2 Stratified Proportional Hazards Models

When adjusting some covariates in a model, we can stratify the corresponding variable

and employ the proportional hazards model for each stratum. In this case, the baseline hazard function is $h_{0i}(t)$, and the effect of other explanatory variables on the hazard function is represented as a proportional hazards model for that stratum. The survival time of the $j$th individual in the $i$th stratum of a population is assumed to follow its own hazard function, which expressed as:

$$h_{ij}\left(t\middle|Z_{ij}(t)\right) = h_{0i}(t)\exp\left(\beta' Z_{ij}(t)\right)$$

where $\beta = \left(\beta_1, \ldots, \beta_p\right)'$ is a parameter of coefficients that estimate the effect of the covariate $Z_{ij}$ of the $j$th individual in the $i$th stratum where $Z_{ij} = \left(Z_{ij1}, \ldots, Z_{ijp}\right)'$ is the vector of covariates for the $j$th individual in the $i$th stratum. Thus, in a stratified proportional hazards model, the regression coefficients are assumed to be the same in each stratum, but the baseline hazard functions for the strata are different and completely irrelevant.

The stratified Cox model performs inference through partial likelihood, but the likelihood is only calculated within stratum $i$ when individual $j$ of stratum $i$ fails. The way the $j$th individual was represented was similar to the way they were in the Cox model, but the difference is that in the stratified proportional hazards model, there is data that differs between strata. The basic data is similar to the Cox model, but the difference is that

there is data from stratum to stratum. In the data, $j$th individual of $i$th stratum has the $(T_{ij}, \delta_{ij}, Z_{ij})$ where $i = 1, \dots, s$ and $j = 1, \dots, n_i$. $t_{i1} < t_{i2} < \cdots < t_{iD}$ denote the ordered observed event times of $i$th stratum and $Z_{(i\ell)k}$ was the $k$th covariate associated with the observation in $i$th stratum that failed at time $t_{i\ell}$. The risk set at time $t_{i\ell}$, $R(t_{i\ell})$, was the set of all observations in $i$th stratum that survived without incident until just before time $t_{i\ell}$.

Thus, the partial likelihood of $i$th stratum was expressed as:

$$L_i(\beta) = \prod_{\ell=1}^{D} \frac{\exp\left[\sum_{k=1}^{p} \beta_k Z_{(i\ell)k}\right]}{\sum_{j \in R(t_{i\ell})} \exp\left[\sum_{k=1}^{p} \beta_k Z_{ijk}\right]}$$

The overall likelihood function was produced by multiplying each stratum's likelihood function together, producing:

$$L(\beta) = \prod_{i=1}^{s} L_i(\beta) = \prod_{i=1}^{s} \prod_{\ell=1}^{D} \frac{\exp\left[\sum_{k=1}^{p} \beta_k Z_{(i\ell)k}\right]}{\sum_{j \in R(t_{i\ell})} \exp\left[\sum_{k=1}^{p} \beta_k Z_{ijk}\right]}$$

$LL(\beta) = \log[L(\beta)]$, so $LL(\beta)$ was expressed as:

$$LL(\beta) = \sum_{i=1}^{s}\sum_{\ell=1}^{D}\sum_{k=1}^{p}\beta_k Z_{(i\ell)k} - \sum_{i=1}^{s}\sum_{\ell=1}^{D}\log\left[\sum_{j\in R(t_{i\ell})}\exp\left(\sum_{k=1}^{p}\beta_k Z_{ijk}\right)\right]$$

The score equation was expressed as:

$$U_h(\beta) = \frac{\partial LL(\beta)}{\partial \beta_h} = \sum_{i=1}^{s}\sum_{\ell=1}^{D} Z_{(i\ell)h} - \sum_{i=1}^{s}\sum_{\ell=1}^{D}\frac{\sum_{j\in R(t_{i\ell})} Z_{ijh}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}{\sum_{j\in R(t_{i\ell})}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}$$

The information matrix is the negative of the matrix of second derivatives of the log likelihood and is given by $I(\beta) = \left[I_{gh}(\beta)\right]_{p\times p}$ with the $(g,h)$th element expressed as:

$$\begin{aligned} & I_{gh}(\beta) \\ &= \sum_{i=1}^{s}\sum_{\ell=1}^{D}\frac{\sum_{j\in R(t_{i\ell})} Z_{ijg}Z_{ijh}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}{\sum_{j\in R(t_{i\ell})}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]} \\ &- \sum_{i=1}^{s}\sum_{\ell=1}^{D}\left[\frac{\sum_{j\in R(t_{i\ell})} Z_{ijg}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}{\sum_{j\in R(t_{i\ell})}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}\right]\left[\frac{\sum_{j\in R(t_{i\ell})} Z_{ijh}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}{\sum_{j\in R(t_{i\ell})}\exp\left[\sum_{k=1}^{p}\beta_k Z_{ijk}\right]}\right] \end{aligned}$$

## 2.3 Frailty Model

The frailty model assumes center effects act proportionally on the baseline risk of failure. However, instead of treating the center effects as parameters, the frailty model treats them as a sample from a member of a family of probability distributions. Frailties are unobservable random effects shared by within-cluster observations. The hazard function of the time to event of the $j$th observation in the $i$th cluster can be expressed by multiplying the proportional hazards model with a random effect as follows:

$$h_{ij}(t|Z_{ij}) = h_0(t)\exp(\beta' Z_{ij})\, u_i$$

where, $h_0(t)$ is the baseline hazard function; $Z_{ij} = (Z_{ij1}, \dots, Z_{ijp})'$ is a covariate vector for the $j$ th observation in the $i$ th cluster; $\beta = (\beta_1, \dots, \beta_p)'$ is defined by the corresponding regression coefficient; $(u_1, \dots, u_s), i = 1, \dots, s,$ is the frailty effect of each cluster, assuming that each $u_i$ is independent of the rest and follows the same distribution. $u_i$ represents how quickly the event occurred for the $i$th cluster relative to the average, which is set to 1. A $u_i$ greater than 1 indicates that the event occurred sooner for the $i$th cluster than the average while a value of less than 1 means that the event occurred later for the $i$th cluster than the average.

We will introduce the gamma frailty model, which assumes that frailty follows a gamma distribution. It is as follows:

$$g(u_i|\theta) = \frac{u_i^{\frac{1}{\theta}-1} \exp\left(-\frac{u_i}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)\theta^{\frac{1}{\theta}}}$$

where $E(u_i) = 1$ and $Var(u_i) = \theta$. The joint survival function using gamma frailty for the $n_i$ observations within the $i$th cluster is as follows:

$$S(t_{i1}, \dots, t_{in_i}) = P(T_{i1} > t_{i1}, \dots, T_{in_i} > t_{in_i}) = \int_0^\infty S(t_{i1}, \dots, t_{in_i}|u_i)\, g(u_i|\theta) du_i$$
$$= \left[1 + \theta \sum_{j=1}^{n_i} H_0(t_{ij}) \exp(\beta' Z_{ij})\right]^{-1/\theta}$$

The association between cluster members as measured by Kendall's $\tau$ is $\theta/(\theta+2)$, and $\theta = 0$ means the case of independence.

Like the model described above, the gamma frailty model performs inference through the partial likelihood. In the data set, $j$th individual of $i$th cluster has the $(T_{ij}, \delta_{ij}, Z_{ij})$,

$i = 1, \ldots, s,\ j = 1, \ldots, n_i.\ D_i = \sum_{j=1}^{n_i} \delta_{ij}$ was the number of events in the $i$th cluster. In proportional hazards models that do not use parametric distributions for baseline hazard functions, EM algorithms should be used to estimate parameters. For this purpose, we used the full likelihood if the frailties were observed, which was expressed as:

$$L_{FULL} = L_1(\theta) \times L_2(\beta, H_0)$$

where:

$$L_1(\theta) = \prod_{i=1}^{s} \frac{u_i^{\frac{1}{\theta}+D_i-1} \exp\left(-\frac{u_i}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right)\theta^{\frac{1}{\theta}}}$$

$$L_2(\beta, H_0) = \prod_{i=1}^{s}\prod_{j=1}^{n_i} \frac{\left(h_0(T_{ij})\exp(\beta' Z_{ij})\right)^{\delta_{ij}}}{\exp\left(H_0(T_{ij})\exp(\beta' Z_{ij})\, u_i\right)}$$

The log-likelihood, *LL(β)*, was expressed as:

$$LL_{FULL} = LL_1(\theta) + LL_2(\beta, H_0)$$

where:

$$LL_1(\theta) = -G\left[\frac{1}{\theta}\log(\theta) + \log\Gamma\left(\frac{1}{\theta}\right)\right] + \sum_{i=1}^{G}\left[\frac{1}{\theta} + D_i - 1\right]\log u_i - \frac{u_i}{\theta}$$

$$LL_2(\beta, H_0) = \sum_{i=1}^{s}\sum_{j=1}^{n_i}\delta_{ij}\left[\beta' Z_{ij} + \log h_0(T_{ij})\right] - u_i H_0(T_{ij})\exp(\beta' Z_{ij})$$

The likelihood function included incomplete data, $u_i$ and $\log u_i$. Thus, the EM algorithm is used for estimation in this process. The expectation–maximization (EM) algorithm is an iterative method to find maximum likelihood estimates of parameters in models, it provides a means of maximizing likelihoods. The EM iteration alternates between performing an expectation (E) step and a maximization (M) step. In the E step, the expected value of $LL_{FULL}$ is computed, given the current estimates of the parameters and the observable data, and in the M step, estimates of the parameters which maximize the expected value of $LL_{FULL}$ found on the E step. The algorithm iterates between these two steps until convergence.

After the parameter were estimated using the EM algorithm, the information matrix was produced using the observable log-likelihood function, which was expressed as:

$$\mathrm{LL}_0(\theta,\beta) = \sum_{i=1}^{s} D_i \log\theta$$
$$- \log\left[\Gamma\left(\frac{1}{\theta}\right)\right] + \log\left[\Gamma\left(\frac{1}{\theta}+D_i\right)\right] - \left(\frac{1}{\theta}+D_i\right)$$
$$\times \log\left[1+\theta\sum_{j=1}^{n_i} H_0(T_{ij})\exp(\beta' Z_{ij})\right] + \sum_{j=1}^{n_i}\delta_{ij}\{\beta' Z_{ij} + \log[h_0(T_{ij})]\}$$

The information matrix $I(\beta) = \left[I_{gh}(\beta)\right]_{p\times p}$ with the $(g,h)$th element was expressed as:

$$I_{gh}(\beta) = \sum_{i=1}^{s} \frac{D_i + \frac{1}{\theta}}{\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(T_{ij})\exp(\beta' Z_{ij})}$$
$$\times \left\{\left[\frac{1}{\theta} + \sum_{j=1}^{n_i} H_0(T_{ij})\exp(\beta' Z_{ij})\right]\right.$$
$$\times \sum_{j=1}^{n_i} H_0(T_{ij}) Z_{ijg} Z_{ijh}\exp(\beta' Z_{ij})$$
$$\left. - \sum_{j=1}^{n_i} H_0(T_{ij}) Z_{ijg}\exp(\beta' Z_{ij}) \sum_{j=1}^{n_i} H_0(T_{ij}) Z_{ijh}\exp(\beta' Z_{ij})\right\}$$

# Chapter 3

# Description of Bias

## 3.1 Marginal Models

The survivor function for the $j$th individual in $i$th stratum is expressed as:

$$S_{ij}(t) = \exp\{-u_i H_0(t) \exp(\beta' Z_{ij})\}$$

where $H_0(t)$ is the baseline cumulative hazard function. This is the survivor function for the $j$th individual in $i$th stratum conditional on the frailty $u_i$, and is termed a conditional model for the survival times. The stratum frailties $u_i$ cannot be observed directly, and so what we observe is the effect that they have on the overall survivor function for the group of strata. The observable survivor function is therefore the individual in stratum functions

averaged over all possible values of $u_i$, that is the expected value of $S_{ij}(t)$ with respect to the frailty distribution. The survivor function that is observed by integrating $S_{ij}(t)$ with respect to the distribution of $u_i$. The resulting unconditional or observable survivor function is represented as:

$$S_{ij}^{*}(t) = \int_0^{\infty} S_{ij}(t) f(u_i)\ du_i = \int_0^{\infty} \exp\{-u_i H_0(t) \exp(\beta' Z_{ij})\} f(u_i)\ du_i$$

where $f(u_i)$ is the density function of the frailty distribution.

For the gamma distribution, each stratum has a distinct frailty value, $u_i$, and that these are independent and identically distributed random variables $u_i, i = 1,2,\dots,n$, where $u_i \sim \Gamma(1/\theta, \theta)$. Now, the observable survivor function is:

$$\begin{aligned} S_{ij}^{*}(t) &= \int_0^{\infty} S_{ij}(t) f(u_i)\ du_i \\ &= \int_0^{\infty} \exp\{-u_i H_0(t) \exp(\beta' Z_{ij})\, H_0(t)\} \frac{u_i^{\frac{1}{\theta}-1} \exp\left(-\frac{u_i}{\theta}\right)}{\Gamma\left(\frac{1}{\theta}\right) \theta^{\frac{1}{\theta}}}\ du_i \\ &= \left\{1 + \theta e^{\beta' x_i} H_0(t)\right\}^{-\frac{1}{\theta}} \end{aligned}$$

For the case, where $Z_{ij}$ is the value of a binary covariate that takes values 0 and 1, the corresponding observable hazard function as follows:

$$h_{ij}^{*}(t) = \frac{h_0(t)\exp(\beta Z_{ij})}{1 + \theta H_0(t)\exp(\beta Z_{ij})}$$

And the observable hazard ratio for $Z = 1$ relative to $Z = 0$ is:

$$\mathrm{HR}^{*}(\mathrm{t}) = \frac{1 + \theta H_0(t)}{1 + \theta H_0(t)\exp(\beta)}\exp(\beta)$$

The ratio of the treatment-specific marginal hazards is not time-invariant; unless $\theta$ or $\beta$ is 0, the marginal hazard ratio is a decreasing function of time. Only at time 0 is the marginal hazard ratio equal to the true conditional hazard ratio $\exp(\beta)$, as $t$ increases and $H_0(t) \to \infty$, the ratio tends to 1.

## 3.2 Omitted Covariates

Following Andersen & Gill (1982), let $(N_1, \dots, N_n), n = 1,2, \dots$ be a multivariate counting process in which $N_i(t)$ records the number of failures. Suppose that $N_i$ has random intensity process of the form:

$$Y_i(t)h_0(t)\exp(\beta Z_i) \quad (i = 1, \dots, n)$$

where $Y_i(t)$, a predictable process, takes the value 1 if the $i$th subject is at risk at time $\mathrm{t}-$ and 0 otherwise, and $Z_i$ is the covariate for the $i$th subject. The regression coefficient $\beta$ was estimated from the partial likelihood, Struthers & Kalbfleisch (1986) showed that partial log-likelihood converges in probability to the concave function $H(\beta)$ which has a unique maximum at $\beta = \beta^*$. Since $\hat{\beta}$ maximizes the partial log-likelihood follows by some convex analysis that $\hat{\beta}$ converges in probability to $\beta^*$.

Thus, the maximum partial likelihood estimator $\hat{\beta}$ is a consistent estimator of $\beta^*$, where $\beta^*$ is the solution to the equation $h(\beta) = 0$. The equation was represented as:

$$h(\beta) = n^{-1}\left[\int_0^{\infty} E\left\{\sum_{i=1}^{n} Z_i Y_i(t) h(t; Z_i)\right\} dt \right.$$

$$\left. - \int_0^{\infty} E\left\{\frac{\sum_{i=1}^{n} Z_i Y_i(t)\exp(\beta Z_i) \sum_{i=1}^{n} Y_i(t) h(t; Z_i)}{\sum_{i=1}^{n} Y_i(t)\exp(\beta Z_i)}\right\} dt\right]$$

It is used to investigate the effects on estimation if covariates are missing from the proportional hazards model.

The stratified Cox model was expressed as:

$$h_{ij}(t) = h_{0i}(t)\exp(\beta Z_{ij}) = h_0(t)\,\gamma_i \exp(\beta Z_{ij}) = h_0(t)\exp(\beta_1 Z_{ij} + \beta_2 \xi_i)$$

When the stratified Cox model is true model, it can be thought of as a missing covariate situation as above, unless stratification is considered. We examine the effect on estimation of $\beta_1$ when the covariates $\xi_i$ are omitted from the model.

In the case of model misspecification, the true model is proportional hazards model with respect to two independent covariates $Z_{ij}$, $\xi_i$ and that the true random intensity process is of the form:

$$Y_i(t)h_0(t)exp(\beta_1 Z_{ij} + \beta_2 \xi_i)$$

where $h_0(t)$ is a baseline hazard function.

We consider the estimation of $\beta_1$ when the model assumed is

$$Y_i(t)h_1(t)\,exp(\beta Z_{ij})$$

For the case $\beta_1 > 0$, $h(\beta)$ was represented as:

$$h(\beta) = \int_0^\infty E\{Z_1 h(t;Z)G(t;Z)\}dt - \int_0^\infty \frac{E\{Z_1 e^{\beta Z_1 + \beta_2 Z_2} G(t;Z)\}}{E\{e^{\beta Z_1 + \beta_2 Z_2} G(t;Z)\}} E\{h(t;Z)G(t;Z)\}dt$$

where :

$$G(t;Z) = C(t;Z_1)\exp\{-H_0(t)\exp(\beta_1 Z_{ij} + \beta_2 \xi_i)\}$$

$$h(t;Z) = h_0(t)\exp(\beta_1 Z_{ij} + \beta_2 \xi_i)$$

Since $h(\beta_1) = 0$ and $dh(\beta)/d\beta < 0$ for all $\beta$, therefore $h(\beta)$ is a monotone decreasing function of $\beta$ which crosses the x axis at $\beta = \beta_1$. It can be shown that $h(\beta^*) > 0$ and therefore $\beta^* < \beta_1$. Likewise, for the case $\beta_1 < 0$, it can be shown that $h(\beta^*) < 0$ and therefore $\beta^* > \beta_1$. The estimator of the regression parameter in the true model is asymptotically biased toward zero.

# Chapter 4

# Simulation

We performed simulations to compare the gamma frailty model and the stratified Cox model according to the difference between functional form of baseline hazards. The simulations were conducted using survival-type datasets.

We provided practical guideline on the performance of the estimates from various situations of the data to investigate in more detail the particular situation in which either the gamma frailty model or the stratified Cox model is preferred. Our interest in this study was to compare estimates of the two models when data is created from the gamma frailty model and when data is created from the stratified Cox proportional hazards model. We compared their bias, MSE, SD, and SE of their results.

## 4.1 Data Generation

We created two types of survival time based on each model to introduce more diversity into the data. The first survival time, $T^*$, was based on the gamma frailty model. The hazard function was assumed to follow the Weibull distribution, the scale parameter was $\lambda$, and the shape parameter was $\alpha$. The random effect term, $u_i$, followed the gamma distribution of $\Gamma(1/\theta, \theta)$. The expectation of the frailty term equals to 1, $E(u_i) = 1$ and the variance is finite, $Var(u_i) = \theta$. We generated hypothetical survival times, $T^*$, for the simulations. The general formula relating the hazard function and the corresponding survival time in the Cox model, the survival time based on the gamma frailty model was expressed as:

$$T^* = \left[ -\frac{\log(s)}{u_i \lambda \exp(\beta Z_{ij})} \right]^{\frac{1}{\alpha}}$$

where $s$ is a random variable with uniform distribution of $\text{unif}(0,1)$.

The second survival time, $T^{**}$, was based on the stratified Cox model. The functional form of the baseline hazards varied from center to center. The hazards function were assumed to follow the Weibull distribution, the scale parameter was $\lambda$, and the shape parameter was $\alpha_i$. The stratified Cox model that eliminated the random effect term from the preceding expression was expressed as:

$$T^{**} = \left[ -\frac{\log(s)}{\lambda \exp(\beta Z_{ij})} \right]^{\frac{1}{\alpha_i}}$$

Furthermore, it was implemented separately between balanced cases and imbalanced cases. In balanced cases, the ratio of treatment group to control group in a stratum was 1:1. In imbalanced cases, the ratio of treatment group to control group in a stratum was depend on the number of strata ($k$); half of the entire stratum maintained a 1:1 ratio, a quarter of the entire stratum assigned 80% of subjects to treatment groups and 20% to control groups, and a quarter of the entire stratum assigned 20% of subjects to treatment groups and 80% to control groups.

Both models included a single covariate: $Z$, where $Z = 1$ for subjects in the treatment group and $Z = 0$ for subjects in the control group.

## 4.2 Simulation Setting

The simulations were run using data generated by both the gamma frailty model and stratified Cox models were run for each of the three models. The treatment effect parameter $\beta$ values for the simulations were 0, -0.3, and -0.5 and the censoring percentages were 25% and 75%. The number of strata ($k$) was 4, 8, 20, and 100 and the total population ($N$) was 400 and 2,000. The number of subjects assigned to each stratum was similar, but not the

same. All analyses used in the simulation were analyzed using R version 4.0.2. Each scenario was repeated 1,000 times independently.

First, we performed the simulation using based on the gamma frailty model. The hazard function was assumed to follow the Weibull distribution, and the scale parameter was set to 1; the shape parameter was set to 2, which was not constant over time and increasing hazard against follow-up time. The gamma distribution was $\Gamma(2, 1/2)$. Thus, to generate data with a true variance of $\theta = 0.5$. The treatment-to-control-group-size ratio in a stratum was 1:1. The second simulation was the same as this one, except that 25% of the strata had only 80% treatment group subjects and 20% control group subjects; 25% of the strata had only 20% treatment group subjects and 80% control group subjects.

Third, we performed the simulation using based on the stratified Cox model. Although the hazard function followed the Weibull distribution, the shape parameter was set to a different value for each stratum such that each stratum had a different baseline. The scale parameters were all set to 1 and the shape parameters were set at values in a range of $0.5-2$. The treatment-to-control-group-size ratio in a stratum was 1:1. The fourth simulation was the same as this one, except that 25% of the strata had only 80% treatment group subjects and 20% control group subjects; 25% of the strata had only 20% treatment group subjects and 80% control group subjects.

The model's performance was evaluated in terms of their mean estimates (Estimate); mean biases (Bias), where the bias was defined as the estimated beta minus the true beta; mean squared errors (MSE), which was defined as $\sum_{i=1}^{n} \frac{(\hat{\beta}-\beta)^2}{n}$; empirical standard

deviation of the parameter estimate (SD); mean estimated standard errors (SE); and empirical power or type I error when $\beta = 0$, which was tested using the Wald statistics. The mean was based on 1,000 replicates.

## 4.3 Results

For the simulations with data generated based on the frailty model and strata were balanced, when the true beta was zero, no models were particularly biased (Table 1). However, for other true beta values, the unadjusted Cox model was more biased than other models. Bias was positively correlated with both the total number of samples and the number of strata (Tables 2 and 3). However, the stratified Cox model was not biased even for simulations using data based on the frailty model. The number of observations in a stratum was negatively correlated with bias. The gamma frailty model was not particularly biased.

For the frailty simulation and strata were imbalanced, the stratified Cox model was not particularly biased, but its standard error was negatively correlated with the total number of samples more strongly than the standard errors for the other models (Tables 4, 5 and 6).

For the simulations with data generated based on the stratified model, when strata were balanced, was presented. None of the models were significantly biased when the true beta was zero (Tables 7), but when it was -0.3 and -0.5, the unadjusted Cox model and frailty

models were biased. (Tables 8 and 9). The models had similar biases regardless of the total number of samples and the number of strata. The stratified Cox model was not particularly biased.

For the stratified simulation and strata were imbalanced, which were similar to those of the frailty simulation and strata were imbalanced. One difference is that unlike frailty simulations, the frailty model had a bias (Tables 10, 11 and 12).

In conclusion, the unadjusted Cox model was biased for all simulations except when the treatment effect was 0, while the stratified Cox model was not biased, no matter which model was correct model. The stratified Cox model fit treatment effect well without significantly affecting the form of the baseline function. The frailty model was not biased for simulations using data generated based on the gamma frailty model but it was for simulations using data based on the stratified Cox model and the size of the bias was negatively correlated with the number of samples in a stratum. The frailty model was biased if the multiplication form of the baseline was not satisfied. However, the stratified model had a larger standard error than the other models when there was an imbalance ratio of groups within the strata. Also, regardless of the imbalance within the strata, the larger the number of centers, the higher the standard error, which showed performance of the low power (Figures 2, 3).

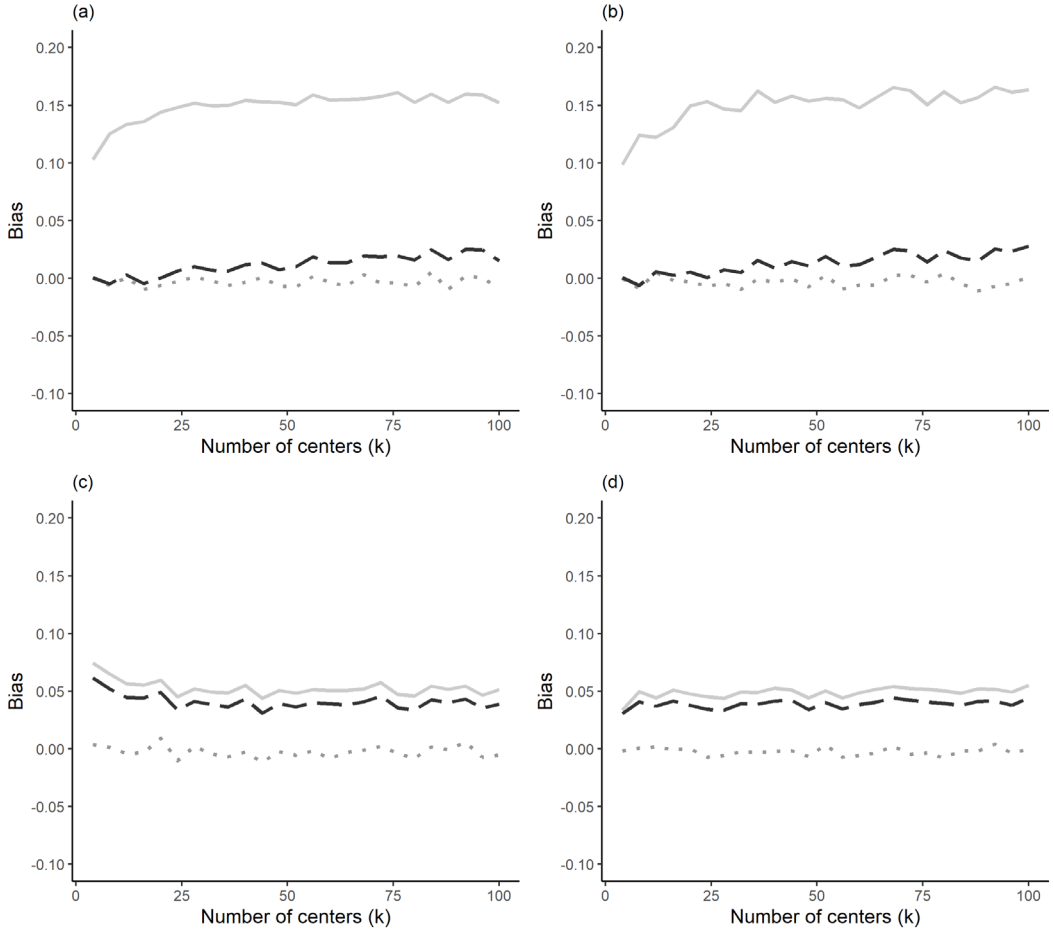


Figure 1. Performance of mean biases with number of centers ($k$) when $N = 400$ and $\beta = -0.5$: (a) frailty simulation and strata were balanced; (b) frailty simulation and strata were imbalanced; (c) stratified simulation and strata were balanced; (d) stratified simulation and strata were imbalanced. Solid line for unadjusted Cox model; dotted line for stratified Cox model; long dashed line for gamma frailty model.

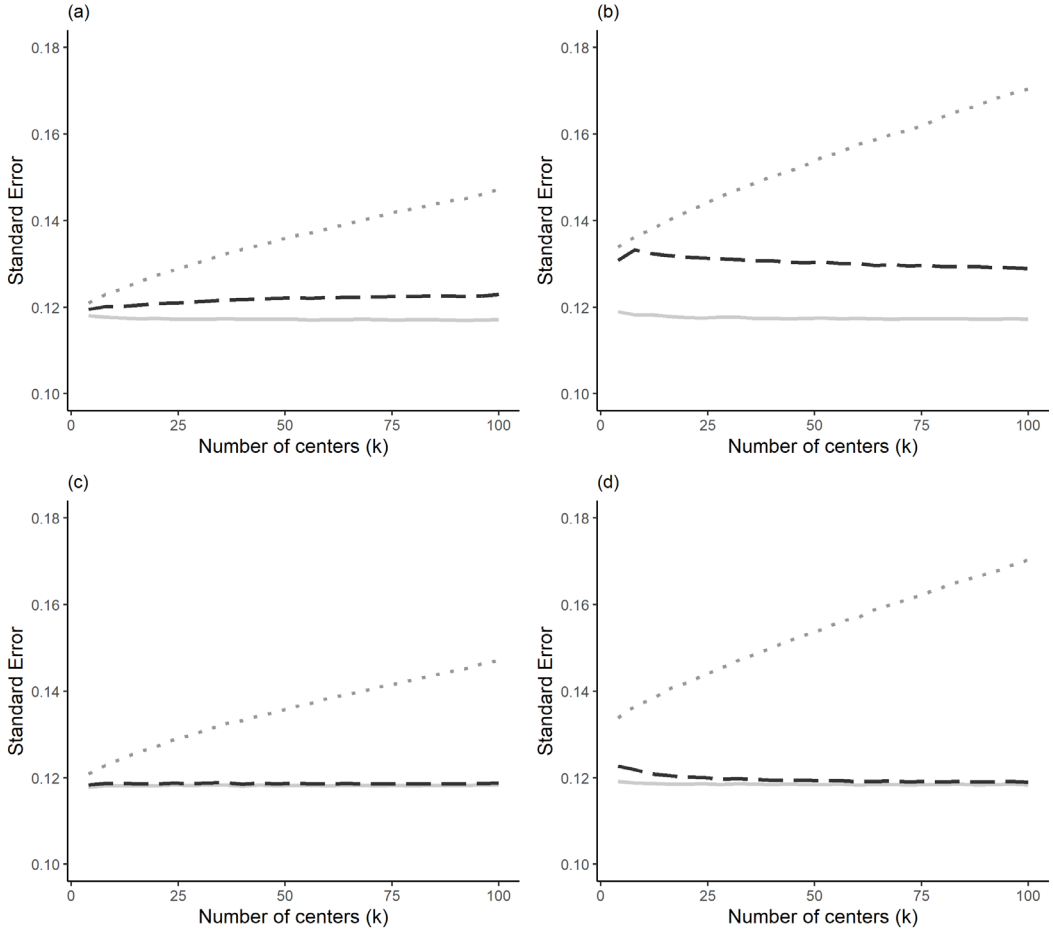


Figure 2. Performance of mean estimated standard errors with number of centers ($k$) when $N = 400$ and $\beta = -0.5$: (a) frailty simulation and strata were balanced; (b) frailty simulation and strata were imbalanced; (c) stratified simulation and strata were balanced; (d) stratified simulation and strata were imbalanced. Solid line for unadjusted Cox model; dotted line for stratified Cox model; long dashed line for gamma frailty model.

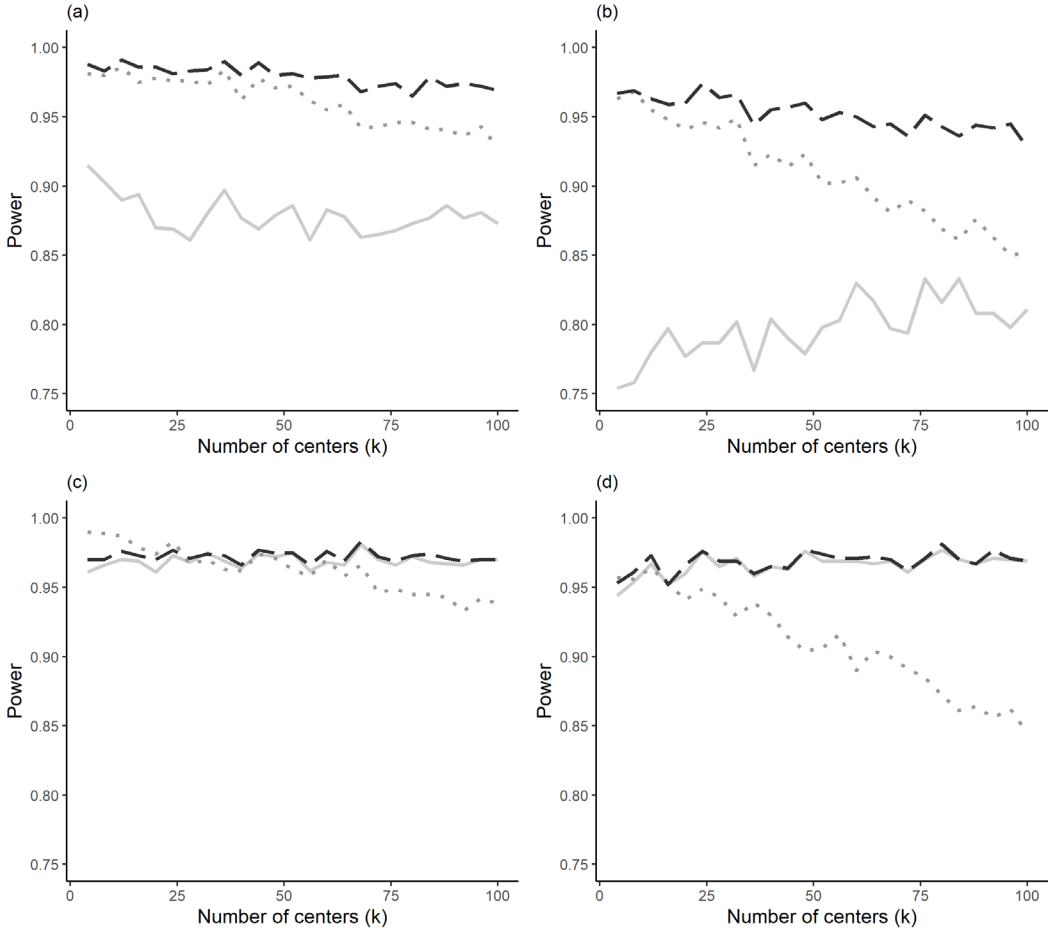


Figure 3. Performance of power with number of centers ($k$) when $N = 400$ and $\beta = -0.5$: (a) frailty simulation and strata were balanced; (b) frailty simulation and strata were imbalanced; (c) stratified simulation and strata were balanced; (d) stratified simulation and strata were imbalanced. Solid line for unadjusted Cox model; dotted line for stratified Cox model; long dashed line for gamma frailty model.

Table 1. Simulation using data based on the gamma frailty model when $\beta = 0,\ \alpha = 2$ and $\theta = 0.5$ and strata were balanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | 0.003 | 0.003 | 0.107 | 0.116 | 0.030 | 0.002 | 0.002 | 0.121 | 0.118 | 0.058 | 0.003 | 0.003 | 0.120 | 0.117 | 0.060 |
| | 8 | 0.002 | 0.002 | 0.105 | 0.116 | 0.030 | 0.002 | 0.002 | 0.123 | 0.120 | 0.054 | 0.003 | 0.003 | 0.120 | 0.117 | 0.056 |
| | 20 | -0.001 | -0.001 | 0.100 | 0.116 | 0.023 | -0.001 | -0.001 | 0.123 | 0.124 | 0.047 | -0.001 | -0.001 | 0.119 | 0.118 | 0.051 |
| | 100 | 0.001 | 0.001 | 0.098 | 0.116 | 0.020 | 0.000 | 0.000 | 0.139 | 0.143 | 0.043 | 0.000 | 0.000 | 0.122 | 0.121 | 0.048 |
| 2000 | 4 | 0.000 | 0.000 | 0.046 | 0.052 | 0.028 | 0.000 | 0.000 | 0.051 | 0.052 | 0.040 | 0.000 | 0.000 | 0.050 | 0.052 | 0.040 |
| | 8 | 0.002 | 0.002 | 0.045 | 0.052 | 0.015 | 0.002 | 0.002 | 0.051 | 0.052 | 0.043 | 0.003 | 0.003 | 0.051 | 0.052 | 0.041 |
| | 20 | 0.002 | 0.002 | 0.045 | 0.052 | 0.025 | 0.002 | 0.002 | 0.053 | 0.053 | 0.057 | 0.002 | 0.002 | 0.053 | 0.052 | 0.057 |
| | 100 | 0.002 | 0.002 | 0.045 | 0.052 | 0.030 | 0.002 | 0.002 | 0.057 | 0.056 | 0.067 | 0.001 | 0.001 | 0.055 | 0.053 | 0.069 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | 0.009 | 0.009 | 0.194 | 0.202 | 0.048 | 0.011 | 0.011 | 0.204 | 0.205 | 0.051 | 0.010 | 0.010 | 0.202 | 0.203 | 0.056 |
| | 8 | -0.005 | -0.005 | 0.201 | 0.203 | 0.050 | -0.003 | -0.003 | 0.210 | 0.209 | 0.052 | -0.005 | -0.005 | 0.208 | 0.204 | 0.053 |
| | 20 | -0.005 | -0.005 | 0.190 | 0.203 | 0.036 | -0.006 | -0.006 | 0.213 | 0.217 | 0.044 | -0.005 | -0.005 | 0.201 | 0.205 | 0.044 |
| | 100 | -0.009 | -0.009 | 0.188 | 0.202 | 0.034 | -0.010 | -0.010 | 0.241 | 0.248 | 0.047 | -0.011 | -0.011 | 0.202 | 0.204 | 0.041 |
| 2000 | 4 | 0.000 | 0.000 | 0.088 | 0.090 | 0.050 | 0.001 | 0.001 | 0.090 | 0.090 | 0.048 | 0.001 | 0.001 | 0.089 | 0.090 | 0.047 |
| | 8 | 0.002 | 0.002 | 0.084 | 0.090 | 0.038 | 0.002 | 0.002 | 0.090 | 0.090 | 0.050 | 0.002 | 0.002 | 0.089 | 0.090 | 0.048 |
| | 20 | 0.003 | 0.003 | 0.087 | 0.090 | 0.040 | 0.002 | 0.002 | 0.092 | 0.092 | 0.043 | 0.002 | 0.002 | 0.091 | 0.090 | 0.044 |
| | 100 | 0.003 | 0.003 | 0.085 | 0.090 | 0.036 | 0.003 | 0.003 | 0.097 | 0.096 | 0.054 | 0.004 | 0.004 | 0.091 | 0.091 | 0.046 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 2. Simulation using data based on the gamma frailty model when $\beta = -0.3$, $\alpha = 2$ and $\theta = 0.5$ and strata were balanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.237 | 0.063 | 0.114 | 0.117 | 0.520 | -0.299 | 0.001 | 0.121 | 0.119 | 0.696 | -0.298 | 0.002 | 0.120 | 0.118 | 0.707 |
| | 8 | -0.225 | 0.075 | 0.108 | 0.117 | 0.471 | -0.304 | -0.004 | 0.125 | 0.121 | 0.702 | -0.304 | -0.004 | 0.123 | 0.118 | 0.719 |
| | 20 | -0.214 | 0.086 | 0.101 | 0.117 | 0.439 | -0.302 | -0.002 | 0.126 | 0.125 | 0.671 | -0.298 | 0.002 | 0.120 | 0.119 | 0.708 |
| | 100 | -0.207 | 0.093 | 0.099 | 0.116 | 0.408 | -0.300 | 0.000 | 0.145 | 0.144 | 0.540 | -0.289 | 0.011 | 0.124 | 0.121 | 0.638 |
| 2000 | 4 | -0.237 | 0.063 | 0.058 | 0.052 | 0.994 | -0.299 | 0.001 | 0.051 | 0.052 | 1.000 | -0.298 | 0.002 | 0.051 | 0.052 | 1.000 |
| | 8 | -0.227 | 0.073 | 0.055 | 0.052 | 0.993 | -0.303 | -0.003 | 0.053 | 0.053 | 1.000 | -0.303 | -0.003 | 0.052 | 0.052 | 1.000 |
| | 20 | -0.212 | 0.088 | 0.049 | 0.052 | 0.990 | -0.301 | -0.001 | 0.055 | 0.053 | 1.000 | -0.301 | -0.001 | 0.054 | 0.052 | 1.000 |
| | 100 | -0.205 | 0.095 | 0.045 | 0.052 | 0.987 | -0.302 | -0.002 | 0.057 | 0.056 | 0.998 | -0.300 | 0.000 | 0.054 | 0.053 | 0.999 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.232 | 0.068 | 0.197 | 0.204 | 0.198 | -0.290 | 0.010 | 0.206 | 0.207 | 0.289 | -0.286 | 0.014 | 0.202 | 0.205 | 0.292 |
| | 8 | -0.233 | 0.067 | 0.199 | 0.203 | 0.191 | -0.312 | -0.012 | 0.208 | 0.211 | 0.298 | -0.299 | 0.001 | 0.205 | 0.205 | 0.299 |
| | 20 | -0.215 | 0.085 | 0.191 | 0.203 | 0.167 | -0.302 | -0.002 | 0.218 | 0.218 | 0.273 | -0.276 | 0.024 | 0.204 | 0.205 | 0.263 |
| | 100 | -0.205 | 0.095 | 0.194 | 0.203 | 0.163 | -0.304 | -0.004 | 0.256 | 0.253 | 0.237 | -0.244 | 0.056 | 0.209 | 0.206 | 0.218 |
| 2000 | 4 | -0.241 | 0.059 | 0.096 | 0.090 | 0.745 | -0.300 | 0.000 | 0.091 | 0.091 | 0.907 | -0.300 | 0.000 | 0.092 | 0.091 | 0.909 |
| | 8 | -0.225 | 0.075 | 0.089 | 0.090 | 0.701 | -0.300 | 0.000 | 0.089 | 0.091 | 0.925 | -0.298 | 0.002 | 0.088 | 0.091 | 0.927 |
| | 20 | -0.209 | 0.091 | 0.091 | 0.090 | 0.639 | -0.299 | 0.001 | 0.094 | 0.092 | 0.893 | -0.294 | 0.006 | 0.094 | 0.091 | 0.889 |
| | 100 | -0.200 | 0.100 | 0.083 | 0.090 | 0.613 | -0.297 | 0.003 | 0.094 | 0.097 | 0.867 | -0.271 | 0.029 | 0.088 | 0.091 | 0.847 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 3. Simulation using data based on the gamma frailty model when $\beta = -0.5$, $\alpha = 2$ and $\theta = 0.5$ and strata were balanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.397 | 0.103 | 0.122 | 0.118 | 0.915 | -0.500 | 0.000 | 0.122 | 0.121 | 0.981 | -0.499 | 0.001 | 0.120 | 0.120 | 0.988 |
| | 8 | -0.375 | 0.125 | 0.115 | 0.118 | 0.903 | -0.506 | -0.006 | 0.128 | 0.123 | 0.980 | -0.505 | -0.005 | 0.124 | 0.120 | 0.983 |
| | 20 | -0.356 | 0.144 | 0.111 | 0.117 | 0.870 | -0.506 | -0.006 | 0.131 | 0.127 | 0.978 | -0.500 | 0.000 | 0.126 | 0.121 | 0.986 |
| | 100 | -0.348 | 0.152 | 0.103 | 0.117 | 0.873 | -0.509 | -0.009 | 0.149 | 0.147 | 0.928 | -0.485 | 0.015 | 0.129 | 0.123 | 0.969 |
| 2000 | 4 | -0.399 | 0.101 | 0.073 | 0.053 | 1.000 | -0.500 | 0.000 | 0.053 | 0.053 | 1.000 | -0.500 | 0.000 | 0.053 | 0.053 | 1.000 |
| | 8 | -0.372 | 0.128 | 0.068 | 0.052 | 1.000 | -0.497 | 0.003 | 0.054 | 0.054 | 1.000 | -0.498 | 0.002 | 0.054 | 0.053 | 1.000 |
| | 20 | -0.356 | 0.144 | 0.057 | 0.052 | 1.000 | -0.500 | 0.000 | 0.052 | 0.054 | 1.000 | -0.502 | -0.002 | 0.052 | 0.053 | 1.000 |
| | 100 | -0.336 | 0.164 | 0.047 | 0.052 | 1.000 | -0.497 | 0.003 | 0.057 | 0.057 | 1.000 | -0.493 | 0.007 | 0.054 | 0.054 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.392 | 0.108 | 0.203 | 0.205 | 0.487 | -0.497 | 0.003 | 0.207 | 0.210 | 0.662 | -0.486 | 0.014 | 0.203 | 0.207 | 0.657 |
| | 8 | -0.387 | 0.113 | 0.206 | 0.205 | 0.489 | -0.516 | -0.016 | 0.217 | 0.214 | 0.677 | -0.494 | 0.006 | 0.210 | 0.208 | 0.663 |
| | 20 | -0.353 | 0.147 | 0.195 | 0.204 | 0.413 | -0.500 | 0.000 | 0.213 | 0.222 | 0.627 | -0.454 | 0.046 | 0.203 | 0.208 | 0.587 |
| | 100 | -0.355 | 0.145 | 0.196 | 0.204 | 0.406 | -0.519 | -0.019 | 0.262 | 0.258 | 0.515 | -0.422 | 0.078 | 0.215 | 0.208 | 0.513 |
| 2000 | 4 | -0.396 | 0.104 | 0.101 | 0.091 | 0.988 | -0.494 | 0.006 | 0.095 | 0.092 | 0.999 | -0.494 | 0.006 | 0.094 | 0.092 | 0.999 |
| | 8 | -0.374 | 0.126 | 0.099 | 0.091 | 0.980 | -0.501 | -0.001 | 0.093 | 0.093 | 1.000 | -0.500 | 0.000 | 0.092 | 0.092 | 1.000 |
| | 20 | -0.358 | 0.142 | 0.094 | 0.091 | 0.980 | -0.504 | -0.004 | 0.092 | 0.094 | 1.000 | -0.496 | 0.004 | 0.090 | 0.092 | 1.000 |
| | 100 | -0.345 | 0.155 | 0.086 | 0.090 | 0.977 | -0.508 | -0.008 | 0.096 | 0.099 | 1.000 | -0.464 | 0.036 | 0.091 | 0.092 | 1.000 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 4. Simulation using data based on the gamma frailty model when $\beta = 0$, $\alpha = 2$ and $\theta = 0.5$ and strata were imbalanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.006 | -0.006 | 0.266 | 0.117 | 0.393 | -0.010 | -0.010 | 0.127 | 0.131 | 0.036 | -0.009 | -0.009 | 0.126 | 0.128 | 0.041 |
| | 8 | 0.001 | 0.001 | 0.213 | 0.117 | 0.273 | -0.005 | -0.005 | 0.134 | 0.133 | 0.044 | -0.005 | -0.005 | 0.131 | 0.131 | 0.043 |
| | 20 | 0.006 | 0.006 | 0.158 | 0.117 | 0.135 | 0.001 | 0.001 | 0.140 | 0.139 | 0.051 | 0.004 | 0.004 | 0.131 | 0.129 | 0.052 |
| | 100 | -0.005 | -0.005 | 0.117 | 0.116 | 0.049 | 0.008 | 0.008 | 0.167 | 0.166 | 0.051 | -0.001 | -0.001 | 0.132 | 0.127 | 0.045 |
| 2000 | 4 | -0.002 | -0.002 | 0.261 | 0.052 | 0.700 | -0.003 | -0.003 | 0.057 | 0.057 | 0.058 | -0.003 | -0.003 | 0.057 | 0.057 | 0.061 |
| | 8 | -0.001 | -0.001 | 0.190 | 0.052 | 0.596 | 0.002 | 0.002 | 0.059 | 0.058 | 0.056 | 0.001 | 0.001 | 0.059 | 0.058 | 0.049 |
| | 20 | -0.006 | -0.006 | 0.129 | 0.052 | 0.413 | -0.001 | -0.001 | 0.057 | 0.058 | 0.051 | -0.001 | -0.001 | 0.056 | 0.058 | 0.041 |
| | 100 | -0.004 | -0.004 | 0.073 | 0.052 | 0.180 | -0.003 | -0.003 | 0.066 | 0.062 | 0.066 | -0.002 | -0.002 | 0.063 | 0.058 | 0.077 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.006 | -0.006 | 0.309 | 0.204 | 0.201 | -0.005 | -0.005 | 0.218 | 0.227 | 0.040 | -0.006 | -0.006 | 0.218 | 0.217 | 0.048 |
| | 8 | 0.016 | 0.016 | 0.268 | 0.203 | 0.138 | 0.003 | 0.003 | 0.237 | 0.231 | 0.048 | 0.006 | 0.006 | 0.228 | 0.219 | 0.047 |
| | 20 | 0.012 | 0.012 | 0.237 | 0.203 | 0.095 | 0.002 | 0.002 | 0.247 | 0.242 | 0.054 | 0.005 | 0.005 | 0.227 | 0.215 | 0.060 |
| | 100 | 0.004 | 0.004 | 0.204 | 0.203 | 0.043 | 0.002 | 0.002 | 0.294 | 0.290 | 0.050 | 0.002 | 0.002 | 0.217 | 0.209 | 0.044 |
| 2000 | 4 | 0.006 | 0.006 | 0.265 | 0.090 | 0.511 | 0.002 | 0.002 | 0.100 | 0.100 | 0.055 | 0.002 | 0.002 | 0.100 | 0.099 | 0.055 |
| | 8 | 0.013 | 0.013 | 0.200 | 0.090 | 0.388 | -0.004 | -0.004 | 0.103 | 0.100 | 0.041 | -0.004 | -0.004 | 0.102 | 0.100 | 0.042 |
| | 20 | -0.004 | -0.004 | 0.142 | 0.090 | 0.207 | -0.001 | -0.001 | 0.098 | 0.101 | 0.041 | 0.000 | 0.000 | 0.096 | 0.099 | 0.041 |
| | 100 | -0.003 | -0.003 | 0.104 | 0.090 | 0.086 | -0.005 | -0.005 | 0.108 | 0.107 | 0.051 | -0.004 | -0.004 | 0.099 | 0.096 | 0.051 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 5. Simulation using data based on the gamma frailty model when $\beta = -0.3$, $\alpha = 2$ and $\theta = 0.5$ and strata were imbalanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.247 | 0.053 | 0.286 | 0.118 | 0.582 | -0.306 | -0.006 | 0.133 | 0.132 | 0.638 | -0.305 | -0.005 | 0.132 | 0.129 | 0.657 |
| | 8 | -0.224 | 0.076 | 0.213 | 0.117 | 0.517 | -0.300 | 0.000 | 0.132 | 0.134 | 0.611 | -0.300 | 0.000 | 0.130 | 0.131 | 0.625 |
| | 20 | -0.217 | 0.083 | 0.158 | 0.117 | 0.485 | -0.303 | -0.003 | 0.142 | 0.140 | 0.579 | -0.301 | -0.001 | 0.132 | 0.130 | 0.616 |
| | 100 | -0.204 | 0.096 | 0.122 | 0.117 | 0.404 | -0.294 | 0.006 | 0.172 | 0.168 | 0.427 | -0.284 | 0.016 | 0.137 | 0.128 | 0.551 |
| 2000 | 4 | -0.243 | 0.057 | 0.253 | 0.052 | 0.833 | -0.299 | 0.001 | 0.057 | 0.058 | 1.000 | -0.300 | 0.000 | 0.057 | 0.058 | 1.000 |
| | 8 | -0.219 | 0.081 | 0.191 | 0.052 | 0.808 | -0.299 | 0.001 | 0.057 | 0.058 | 0.999 | -0.299 | 0.001 | 0.057 | 0.059 | 1.000 |
| | 20 | -0.217 | 0.083 | 0.129 | 0.052 | 0.830 | -0.301 | -0.001 | 0.056 | 0.059 | 1.000 | -0.302 | -0.002 | 0.055 | 0.058 | 1.000 |
| | 100 | -0.206 | 0.094 | 0.077 | 0.052 | 0.913 | -0.302 | -0.002 | 0.063 | 0.062 | 0.997 | -0.300 | 0.000 | 0.060 | 0.058 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.234 | 0.066 | 0.325 | 0.205 | 0.326 | -0.308 | -0.008 | 0.227 | 0.230 | 0.259 | -0.301 | -0.001 | 0.229 | 0.220 | 0.264 |
| | 8 | -0.208 | 0.092 | 0.269 | 0.205 | 0.253 | -0.284 | 0.016 | 0.237 | 0.234 | 0.232 | -0.272 | 0.028 | 0.227 | 0.221 | 0.226 |
| | 20 | -0.216 | 0.084 | 0.227 | 0.204 | 0.202 | -0.302 | -0.002 | 0.248 | 0.244 | 0.251 | -0.274 | 0.026 | 0.222 | 0.217 | 0.235 |
| | 100 | -0.195 | 0.105 | 0.208 | 0.204 | 0.160 | -0.304 | -0.004 | 0.286 | 0.294 | 0.162 | -0.238 | 0.062 | 0.220 | 0.210 | 0.190 |
| 2000 | 4 | -0.227 | 0.073 | 0.270 | 0.091 | 0.667 | -0.296 | 0.004 | 0.095 | 0.101 | 0.849 | -0.295 | 0.005 | 0.095 | 0.100 | 0.845 |
| | 8 | -0.214 | 0.086 | 0.204 | 0.091 | 0.628 | -0.296 | 0.004 | 0.104 | 0.101 | 0.832 | -0.295 | 0.005 | 0.104 | 0.101 | 0.829 |
| | 20 | -0.215 | 0.085 | 0.150 | 0.090 | 0.591 | -0.307 | -0.007 | 0.102 | 0.102 | 0.857 | -0.302 | -0.002 | 0.100 | 0.100 | 0.853 |
| | 100 | -0.206 | 0.094 | 0.102 | 0.090 | 0.605 | -0.300 | 0.000 | 0.106 | 0.108 | 0.801 | -0.274 | 0.026 | 0.094 | 0.096 | 0.803 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 6. Simulation using data based on the gamma frailty model when $\beta = -0.5$, $\alpha = 2$ and $\theta = 0.5$ and strata were imbalanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.401 | 0.099 | 0.289 | 0.119 | 0.754 | -0.500 | 0.000 | 0.132 | 0.134 | 0.963 | -0.499 | 0.001 | 0.132 | 0.131 | 0.967 |
| | 8 | -0.376 | 0.124 | 0.226 | 0.118 | 0.758 | -0.509 | -0.009 | 0.134 | 0.136 | 0.968 | -0.506 | -0.006 | 0.132 | 0.133 | 0.969 |
| | 20 | -0.350 | 0.150 | 0.162 | 0.118 | 0.777 | -0.503 | -0.003 | 0.142 | 0.142 | 0.941 | -0.495 | 0.005 | 0.132 | 0.132 | 0.960 |
| | 100 | -0.336 | 0.164 | 0.121 | 0.117 | 0.811 | -0.499 | 0.001 | 0.168 | 0.170 | 0.852 | -0.472 | 0.028 | 0.141 | 0.129 | 0.928 |
| 2000 | 4 | -0.399 | 0.101 | 0.262 | 0.053 | 0.901 | -0.499 | 0.001 | 0.058 | 0.059 | 1.000 | -0.500 | 0.000 | 0.058 | 0.058 | 1.000 |
| | 8 | -0.368 | 0.132 | 0.193 | 0.053 | 0.917 | -0.499 | 0.001 | 0.057 | 0.059 | 1.000 | -0.500 | 0.000 | 0.056 | 0.059 | 1.000 |
| | 20 | -0.356 | 0.144 | 0.131 | 0.052 | 0.967 | -0.500 | 0.000 | 0.060 | 0.060 | 1.000 | -0.501 | -0.001 | 0.059 | 0.059 | 1.000 |
| | 100 | -0.340 | 0.160 | 0.079 | 0.052 | 0.999 | -0.499 | 0.001 | 0.062 | 0.063 | 1.000 | -0.497 | 0.003 | 0.059 | 0.059 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.401 | 0.099 | 0.336 | 0.207 | 0.521 | -0.500 | 0.000 | 0.245 | 0.233 | 0.588 | -0.489 | 0.011 | 0.242 | 0.222 | 0.605 |
| | 8 | -0.397 | 0.103 | 0.285 | 0.207 | 0.488 | -0.508 | -0.008 | 0.239 | 0.237 | 0.564 | -0.489 | 0.011 | 0.234 | 0.224 | 0.572 |
| | 20 | -0.346 | 0.154 | 0.230 | 0.205 | 0.387 | -0.488 | 0.012 | 0.242 | 0.248 | 0.517 | -0.442 | 0.058 | 0.225 | 0.218 | 0.500 |
| | 100 | -0.342 | 0.158 | 0.208 | 0.204 | 0.396 | -0.520 | -0.020 | 0.298 | 0.299 | 0.408 | -0.414 | 0.086 | 0.224 | 0.211 | 0.475 |
| 2000 | 4 | -0.399 | 0.101 | 0.271 | 0.092 | 0.811 | -0.497 | 0.003 | 0.106 | 0.102 | 0.997 | -0.496 | 0.004 | 0.105 | 0.101 | 0.997 |
| | 8 | -0.385 | 0.115 | 0.204 | 0.091 | 0.848 | -0.503 | -0.003 | 0.105 | 0.103 | 1.000 | -0.500 | 0.000 | 0.103 | 0.103 | 1.000 |
| | 20 | -0.357 | 0.143 | 0.145 | 0.091 | 0.888 | -0.502 | -0.002 | 0.104 | 0.104 | 1.000 | -0.494 | 0.006 | 0.101 | 0.101 | 1.000 |
| | 100 | -0.341 | 0.159 | 0.107 | 0.090 | 0.937 | -0.501 | -0.001 | 0.110 | 0.110 | 0.998 | -0.460 | 0.040 | 0.101 | 0.097 | 0.997 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 7. Simulation using data based on the stratified Cox model when $\beta = 0$ and $\alpha_i = 0.5\sim2$ and strata were balanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | 0.005 | 0.005 | 0.116 | 0.116 | 0.046 | 0.005 | 0.005 | 0.119 | 0.118 | 0.050 | 0.005 | 0.005 | 0.115 | 0.116 | 0.046 |
| | 8 | -0.003 | -0.003 | 0.114 | 0.116 | 0.055 | -0.003 | -0.003 | 0.118 | 0.120 | 0.053 | -0.003 | -0.003 | 0.114 | 0.116 | 0.053 |
| | 20 | 0.006 | 0.006 | 0.116 | 0.116 | 0.044 | 0.002 | 0.002 | 0.127 | 0.124 | 0.057 | 0.005 | 0.005 | 0.116 | 0.116 | 0.045 |
| | 100 | 0.001 | 0.001 | 0.111 | 0.116 | 0.039 | -0.001 | -0.001 | 0.137 | 0.143 | 0.038 | 0.001 | 0.001 | 0.113 | 0.116 | 0.041 |
| 2000 | 4 | -0.001 | -0.001 | 0.053 | 0.052 | 0.058 | -0.001 | -0.001 | 0.054 | 0.052 | 0.061 | -0.001 | -0.001 | 0.052 | 0.052 | 0.054 |
| | 8 | -0.001 | -0.001 | 0.050 | 0.052 | 0.046 | 0.000 | 0.000 | 0.052 | 0.052 | 0.046 | 0.000 | 0.000 | 0.049 | 0.052 | 0.041 |
| | 20 | 0.001 | 0.001 | 0.053 | 0.052 | 0.066 | 0.000 | 0.000 | 0.054 | 0.053 | 0.056 | 0.001 | 0.001 | 0.053 | 0.052 | 0.059 |
| | 100 | -0.001 | -0.001 | 0.050 | 0.052 | 0.045 | 0.000 | 0.000 | 0.053 | 0.056 | 0.043 | -0.001 | -0.001 | 0.050 | 0.052 | 0.047 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.009 | -0.009 | 0.203 | 0.203 | 0.048 | -0.010 | -0.010 | 0.206 | 0.206 | 0.050 | -0.009 | -0.009 | 0.204 | 0.203 | 0.045 |
| | 8 | -0.004 | -0.004 | 0.207 | 0.202 | 0.055 | -0.005 | -0.005 | 0.213 | 0.209 | 0.052 | -0.003 | -0.003 | 0.208 | 0.202 | 0.053 |
| | 20 | -0.002 | -0.002 | 0.205 | 0.202 | 0.049 | -0.001 | -0.001 | 0.222 | 0.216 | 0.047 | -0.002 | -0.002 | 0.207 | 0.202 | 0.050 |
| | 100 | 0.002 | 0.002 | 0.209 | 0.202 | 0.056 | 0.008 | 0.008 | 0.260 | 0.250 | 0.049 | 0.002 | 0.002 | 0.212 | 0.203 | 0.056 |
| 2000 | 4 | -0.004 | -0.004 | 0.093 | 0.090 | 0.065 | -0.004 | -0.004 | 0.094 | 0.090 | 0.067 | -0.004 | -0.004 | 0.093 | 0.090 | 0.059 |
| | 8 | 0.004 | 0.004 | 0.089 | 0.090 | 0.048 | 0.003 | 0.003 | 0.090 | 0.090 | 0.052 | 0.004 | 0.004 | 0.089 | 0.090 | 0.044 |
| | 20 | 0.000 | 0.000 | 0.087 | 0.090 | 0.039 | 0.000 | 0.000 | 0.088 | 0.092 | 0.039 | 0.000 | 0.000 | 0.087 | 0.090 | 0.038 |
| | 100 | -0.007 | -0.007 | 0.088 | 0.090 | 0.041 | -0.006 | -0.006 | 0.094 | 0.096 | 0.039 | -0.007 | -0.007 | 0.089 | 0.090 | 0.041 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 8. Simulation using data based on the stratified Cox model when $\beta = -0.3$ and $\alpha_i = 0.5{\sim}2$ and strata were balanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.261 | 0.039 | 0.110 | 0.117 | 0.619 | -0.301 | -0.001 | 0.115 | 0.119 | 0.724 | -0.266 | 0.034 | 0.109 | 0.117 | 0.634 |
| | 8 | -0.262 | 0.038 | 0.115 | 0.117 | 0.613 | -0.296 | 0.004 | 0.122 | 0.121 | 0.675 | -0.268 | 0.032 | 0.116 | 0.117 | 0.626 |
| | 20 | -0.268 | 0.032 | 0.117 | 0.117 | 0.628 | -0.296 | 0.004 | 0.124 | 0.125 | 0.652 | -0.272 | 0.028 | 0.118 | 0.117 | 0.635 |
| | 100 | -0.274 | 0.026 | 0.115 | 0.117 | 0.667 | -0.304 | -0.004 | 0.145 | 0.145 | 0.564 | -0.281 | 0.019 | 0.117 | 0.117 | 0.683 |
| 2000 | 4 | -0.261 | 0.039 | 0.050 | 0.052 | 1.000 | -0.301 | -0.001 | 0.051 | 0.052 | 1.000 | -0.267 | 0.033 | 0.049 | 0.052 | 1.000 |
| | 8 | -0.265 | 0.035 | 0.053 | 0.052 | 1.000 | -0.301 | -0.001 | 0.054 | 0.053 | 1.000 | -0.271 | 0.029 | 0.052 | 0.052 | 0.999 |
| | 20 | -0.270 | 0.030 | 0.054 | 0.052 | 1.000 | -0.303 | -0.003 | 0.055 | 0.053 | 1.000 | -0.276 | 0.024 | 0.053 | 0.052 | 1.000 |
| | 100 | -0.271 | 0.029 | 0.050 | 0.052 | 1.000 | -0.302 | -0.002 | 0.054 | 0.056 | 1.000 | -0.277 | 0.023 | 0.051 | 0.052 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.258 | 0.042 | 0.204 | 0.204 | 0.246 | -0.296 | 0.004 | 0.204 | 0.208 | 0.294 | -0.261 | 0.039 | 0.203 | 0.204 | 0.247 |
| | 8 | -0.267 | 0.033 | 0.196 | 0.203 | 0.256 | -0.301 | -0.001 | 0.202 | 0.210 | 0.299 | -0.270 | 0.030 | 0.197 | 0.203 | 0.261 |
| | 20 | -0.268 | 0.032 | 0.210 | 0.204 | 0.263 | -0.304 | -0.004 | 0.227 | 0.219 | 0.285 | -0.271 | 0.029 | 0.211 | 0.204 | 0.267 |
| | 100 | -0.278 | 0.022 | 0.208 | 0.204 | 0.275 | -0.308 | -0.008 | 0.265 | 0.253 | 0.243 | -0.283 | 0.017 | 0.212 | 0.204 | 0.277 |
| 2000 | 4 | -0.265 | 0.035 | 0.092 | 0.090 | 0.835 | -0.303 | -0.003 | 0.094 | 0.091 | 0.912 | -0.269 | 0.031 | 0.092 | 0.090 | 0.844 |
| | 8 | -0.266 | 0.034 | 0.090 | 0.090 | 0.833 | -0.301 | -0.001 | 0.092 | 0.091 | 0.907 | -0.270 | 0.030 | 0.090 | 0.090 | 0.839 |
| | 20 | -0.266 | 0.034 | 0.092 | 0.090 | 0.834 | -0.297 | 0.003 | 0.094 | 0.092 | 0.898 | -0.269 | 0.031 | 0.092 | 0.090 | 0.842 |
| | 100 | -0.267 | 0.033 | 0.093 | 0.090 | 0.839 | -0.297 | 0.003 | 0.100 | 0.097 | 0.869 | -0.269 | 0.031 | 0.094 | 0.090 | 0.840 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 9. Simulation using data based on the stratified Cox model when $\beta = -0.5$ and $\alpha_i = 0.5 \sim 2$ and strata were balanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.426 | 0.074 | 0.116 | 0.118 | 0.961 | -0.496 | 0.004 | 0.119 | 0.121 | 0.990 | -0.439 | 0.061 | 0.113 | 0.118 | 0.970 |
| | 8 | -0.435 | 0.065 | 0.116 | 0.118 | 0.966 | -0.499 | 0.001 | 0.122 | 0.123 | 0.989 | -0.448 | 0.052 | 0.117 | 0.119 | 0.970 |
| | 20 | -0.441 | 0.059 | 0.121 | 0.118 | 0.961 | -0.491 | 0.009 | 0.128 | 0.127 | 0.974 | -0.451 | 0.049 | 0.122 | 0.119 | 0.970 |
| | 100 | -0.449 | 0.051 | 0.117 | 0.118 | 0.970 | -0.505 | -0.005 | 0.143 | 0.147 | 0.938 | -0.461 | 0.039 | 0.121 | 0.119 | 0.970 |
| 2000 | 4 | -0.429 | 0.071 | 0.051 | 0.053 | 1.000 | -0.501 | -0.001 | 0.053 | 0.053 | 1.000 | -0.442 | 0.058 | 0.050 | 0.053 | 1.000 |
| | 8 | -0.438 | 0.062 | 0.053 | 0.053 | 1.000 | -0.501 | -0.001 | 0.054 | 0.054 | 1.000 | -0.452 | 0.048 | 0.051 | 0.053 | 1.000 |
| | 20 | -0.441 | 0.059 | 0.051 | 0.053 | 1.000 | -0.498 | 0.002 | 0.053 | 0.054 | 1.000 | -0.454 | 0.046 | 0.050 | 0.053 | 1.000 |
| | 100 | -0.449 | 0.051 | 0.052 | 0.053 | 1.000 | -0.502 | -0.002 | 0.057 | 0.057 | 1.000 | -0.461 | 0.039 | 0.053 | 0.053 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.432 | 0.068 | 0.207 | 0.205 | 0.548 | -0.501 | -0.001 | 0.214 | 0.211 | 0.659 | -0.439 | 0.061 | 0.207 | 0.206 | 0.564 |
| | 8 | -0.438 | 0.062 | 0.197 | 0.205 | 0.580 | -0.501 | -0.001 | 0.205 | 0.214 | 0.663 | -0.445 | 0.055 | 0.197 | 0.206 | 0.588 |
| | 20 | -0.437 | 0.063 | 0.203 | 0.205 | 0.584 | -0.496 | 0.004 | 0.220 | 0.222 | 0.621 | -0.445 | 0.055 | 0.205 | 0.206 | 0.598 |
| | 100 | -0.443 | 0.057 | 0.203 | 0.206 | 0.564 | -0.500 | 0.000 | 0.253 | 0.257 | 0.502 | -0.450 | 0.050 | 0.207 | 0.206 | 0.574 |
| 2000 | 4 | -0.432 | 0.068 | 0.090 | 0.091 | 0.997 | -0.504 | -0.004 | 0.093 | 0.093 | 1.000 | -0.444 | 0.056 | 0.090 | 0.091 | 0.999 |
| | 8 | -0.439 | 0.061 | 0.090 | 0.091 | 0.998 | -0.502 | -0.002 | 0.092 | 0.093 | 0.999 | -0.450 | 0.050 | 0.090 | 0.091 | 0.999 |
| | 20 | -0.446 | 0.054 | 0.091 | 0.091 | 0.997 | -0.502 | -0.002 | 0.093 | 0.094 | 0.998 | -0.454 | 0.046 | 0.091 | 0.091 | 0.998 |
| | 100 | -0.447 | 0.053 | 0.090 | 0.091 | 0.999 | -0.503 | -0.003 | 0.097 | 0.099 | 0.999 | -0.452 | 0.048 | 0.091 | 0.091 | 0.999 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 10. Simulation using data based on the stratified Cox model when $\beta = 0$ and $\alpha_i = 0.5\sim2$ and strata were imbalanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error | Estimate | Bias | SD | SE | Type I error |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | 0.007 | 0.007 | 0.129 | 0.117 | 0.071 | 0.005 | 0.005 | 0.134 | 0.131 | 0.057 | 0.006 | 0.006 | 0.127 | 0.118 | 0.064 |
| | 8 | 0.008 | 0.008 | 0.126 | 0.117 | 0.074 | 0.009 | 0.009 | 0.139 | 0.133 | 0.062 | 0.008 | 0.008 | 0.126 | 0.118 | 0.072 |
| | 20 | 0.006 | 0.006 | 0.117 | 0.116 | 0.045 | 0.009 | 0.009 | 0.138 | 0.139 | 0.047 | 0.006 | 0.006 | 0.118 | 0.117 | 0.047 |
| | 100 | 0.002 | 0.002 | 0.125 | 0.116 | 0.070 | 0.002 | 0.002 | 0.174 | 0.166 | 0.059 | 0.002 | 0.002 | 0.127 | 0.117 | 0.073 |
| 2000 | 4 | 0.003 | 0.003 | 0.071 | 0.052 | 0.159 | -0.001 | -0.001 | 0.058 | 0.057 | 0.058 | 0.000 | 0.000 | 0.058 | 0.055 | 0.060 |
| | 8 | -0.001 | -0.001 | 0.063 | 0.052 | 0.116 | 0.000 | 0.000 | 0.056 | 0.058 | 0.042 | 0.000 | 0.000 | 0.055 | 0.054 | 0.050 |
| | 20 | -0.001 | -0.001 | 0.055 | 0.052 | 0.065 | 0.000 | 0.000 | 0.058 | 0.058 | 0.038 | -0.001 | -0.001 | 0.053 | 0.053 | 0.046 |
| | 100 | -0.003 | -0.003 | 0.052 | 0.052 | 0.055 | -0.001 | -0.001 | 0.063 | 0.062 | 0.053 | -0.003 | -0.003 | 0.053 | 0.052 | 0.050 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | 0.006 | 0.006 | 0.214 | 0.204 | 0.063 | 0.003 | 0.003 | 0.235 | 0.228 | 0.055 | 0.005 | 0.005 | 0.215 | 0.205 | 0.059 |
| | 8 | 0.003 | 0.003 | 0.206 | 0.203 | 0.050 | 0.001 | 0.001 | 0.237 | 0.232 | 0.055 | 0.004 | 0.004 | 0.208 | 0.204 | 0.048 |
| | 20 | 0.006 | 0.006 | 0.199 | 0.203 | 0.045 | 0.007 | 0.007 | 0.242 | 0.242 | 0.049 | 0.007 | 0.007 | 0.201 | 0.203 | 0.048 |
| | 100 | 0.013 | 0.013 | 0.206 | 0.203 | 0.055 | 0.009 | 0.009 | 0.300 | 0.291 | 0.054 | 0.013 | 0.013 | 0.210 | 0.203 | 0.058 |
| 2000 | 4 | -0.004 | -0.004 | 0.102 | 0.090 | 0.085 | -0.003 | -0.003 | 0.100 | 0.099 | 0.044 | -0.004 | -0.004 | 0.098 | 0.092 | 0.062 |
| | 8 | 0.000 | 0.000 | 0.093 | 0.090 | 0.062 | 0.002 | 0.002 | 0.099 | 0.100 | 0.049 | 0.001 | 0.001 | 0.093 | 0.091 | 0.055 |
| | 20 | 0.002 | 0.002 | 0.096 | 0.090 | 0.069 | 0.004 | 0.004 | 0.105 | 0.101 | 0.060 | 0.002 | 0.002 | 0.096 | 0.090 | 0.061 |
| | 100 | -0.003 | -0.003 | 0.094 | 0.090 | 0.064 | -0.004 | -0.004 | 0.111 | 0.107 | 0.054 | -0.004 | -0.004 | 0.094 | 0.090 | 0.064 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 11. Simulation using data based on the stratified Cox model when $\beta = -0.3$ and $\alpha_i = 0.5{\sim}2$ and strata were imbalanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.283 | 0.017 | 0.136 | 0.118 | 0.656 | -0.303 | -0.003 | 0.135 | 0.132 | 0.639 | -0.285 | 0.015 | 0.132 | 0.120 | 0.652 |
| | 8 | -0.266 | 0.034 | 0.127 | 0.117 | 0.621 | -0.297 | 0.003 | 0.128 | 0.134 | 0.616 | -0.270 | 0.030 | 0.122 | 0.119 | 0.618 |
| | 20 | -0.270 | 0.030 | 0.123 | 0.117 | 0.628 | -0.300 | 0.000 | 0.143 | 0.140 | 0.567 | -0.275 | 0.025 | 0.124 | 0.118 | 0.629 |
| | 100 | -0.270 | 0.030 | 0.122 | 0.117 | 0.641 | -0.306 | -0.006 | 0.172 | 0.168 | 0.467 | -0.276 | 0.024 | 0.123 | 0.117 | 0.649 |
| 2000 | 4 | -0.282 | 0.018 | 0.086 | 0.052 | 0.981 | -0.300 | 0.000 | 0.058 | 0.058 | 0.998 | -0.283 | 0.017 | 0.059 | 0.056 | 0.997 |
| | 8 | -0.273 | 0.027 | 0.071 | 0.052 | 0.994 | -0.302 | -0.002 | 0.057 | 0.058 | 1.000 | -0.280 | 0.020 | 0.057 | 0.056 | 0.999 |
| | 20 | -0.271 | 0.029 | 0.061 | 0.052 | 0.999 | -0.300 | 0.000 | 0.058 | 0.059 | 1.000 | -0.276 | 0.024 | 0.056 | 0.055 | 0.999 |
| | 100 | -0.271 | 0.029 | 0.052 | 0.052 | 1.000 | -0.301 | -0.001 | 0.061 | 0.062 | 0.999 | -0.276 | 0.024 | 0.052 | 0.053 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.275 | 0.025 | 0.217 | 0.205 | 0.281 | -0.293 | 0.007 | 0.231 | 0.229 | 0.262 | -0.276 | 0.024 | 0.216 | 0.207 | 0.272 |
| | 8 | -0.282 | 0.018 | 0.215 | 0.204 | 0.283 | -0.305 | -0.005 | 0.241 | 0.234 | 0.269 | -0.286 | 0.014 | 0.216 | 0.206 | 0.281 |
| | 20 | -0.293 | 0.007 | 0.209 | 0.204 | 0.298 | -0.316 | -0.016 | 0.248 | 0.245 | 0.257 | -0.295 | 0.005 | 0.210 | 0.205 | 0.300 |
| | 100 | -0.264 | 0.036 | 0.210 | 0.204 | 0.250 | -0.295 | 0.005 | 0.299 | 0.294 | 0.166 | -0.267 | 0.033 | 0.214 | 0.205 | 0.254 |
| 2000 | 4 | -0.278 | 0.022 | 0.112 | 0.091 | 0.805 | -0.300 | 0.000 | 0.104 | 0.100 | 0.852 | -0.280 | 0.020 | 0.103 | 0.094 | 0.810 |
| | 8 | -0.276 | 0.024 | 0.103 | 0.091 | 0.824 | -0.301 | -0.001 | 0.102 | 0.101 | 0.832 | -0.279 | 0.021 | 0.098 | 0.093 | 0.825 |
| | 20 | -0.269 | 0.031 | 0.097 | 0.090 | 0.838 | -0.295 | 0.005 | 0.105 | 0.102 | 0.827 | -0.271 | 0.029 | 0.096 | 0.092 | 0.820 |
| | 100 | -0.265 | 0.035 | 0.089 | 0.090 | 0.831 | -0.297 | 0.003 | 0.109 | 0.108 | 0.773 | -0.267 | 0.033 | 0.090 | 0.090 | 0.830 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Table 12. Simulation using data based on the stratified Cox model when $\beta = -0.5$ and $\alpha_i = 0.5 \sim 2$ and strata were imbalanced.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.466 | 0.034 | 0.146 | 0.119 | 0.944 | -0.502 | -0.002 | 0.137 | 0.134 | 0.957 | -0.469 | 0.031 | 0.134 | 0.123 | 0.953 |
| | 8 | -0.450 | 0.050 | 0.132 | 0.119 | 0.954 | -0.499 | 0.001 | 0.138 | 0.137 | 0.956 | -0.459 | 0.041 | 0.128 | 0.122 | 0.961 |
| | 20 | -0.453 | 0.047 | 0.125 | 0.118 | 0.960 | -0.500 | 0.000 | 0.142 | 0.142 | 0.941 | -0.462 | 0.038 | 0.125 | 0.120 | 0.966 |
| | 100 | -0.445 | 0.055 | 0.123 | 0.118 | 0.969 | -0.500 | 0.000 | 0.170 | 0.170 | 0.845 | -0.456 | 0.044 | 0.126 | 0.119 | 0.969 |
| 2000 | 4 | -0.461 | 0.039 | 0.097 | 0.053 | 1.000 | -0.501 | -0.001 | 0.058 | 0.059 | 1.000 | -0.471 | 0.029 | 0.062 | 0.057 | 1.000 |
| | 8 | -0.453 | 0.047 | 0.080 | 0.053 | 1.000 | -0.501 | -0.001 | 0.060 | 0.059 | 1.000 | -0.464 | 0.036 | 0.059 | 0.058 | 1.000 |
| | 20 | -0.450 | 0.050 | 0.065 | 0.053 | 1.000 | -0.501 | -0.001 | 0.058 | 0.060 | 1.000 | -0.461 | 0.039 | 0.056 | 0.056 | 1.000 |
| | 100 | -0.444 | 0.056 | 0.055 | 0.053 | 1.000 | -0.497 | 0.003 | 0.062 | 0.063 | 1.000 | -0.454 | 0.046 | 0.055 | 0.053 | 1.000 |
| **Censoring = 75%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.470 | 0.030 | 0.227 | 0.207 | 0.607 | -0.511 | -0.011 | 0.237 | 0.233 | 0.594 | -0.474 | 0.026 | 0.226 | 0.210 | 0.603 |
| | 8 | -0.457 | 0.043 | 0.213 | 0.208 | 0.607 | -0.504 | -0.004 | 0.231 | 0.238 | 0.574 | -0.461 | 0.039 | 0.211 | 0.210 | 0.601 |
| | 20 | -0.448 | 0.052 | 0.210 | 0.207 | 0.589 | -0.505 | -0.005 | 0.253 | 0.248 | 0.548 | -0.454 | 0.046 | 0.212 | 0.208 | 0.588 |
| | 100 | -0.445 | 0.055 | 0.213 | 0.206 | 0.592 | -0.506 | -0.006 | 0.306 | 0.299 | 0.394 | -0.453 | 0.047 | 0.217 | 0.207 | 0.594 |
| 2000 | 4 | -0.463 | 0.037 | 0.124 | 0.092 | 0.983 | -0.498 | 0.002 | 0.106 | 0.102 | 1.000 | -0.466 | 0.034 | 0.108 | 0.096 | 0.996 |
| | 8 | -0.453 | 0.047 | 0.109 | 0.092 | 0.996 | -0.496 | 0.004 | 0.101 | 0.102 | 1.000 | -0.459 | 0.041 | 0.099 | 0.095 | 0.998 |
| | 20 | -0.447 | 0.053 | 0.099 | 0.091 | 0.991 | -0.501 | -0.001 | 0.101 | 0.104 | 0.996 | -0.454 | 0.046 | 0.095 | 0.093 | 0.992 |
| | 100 | -0.447 | 0.053 | 0.092 | 0.091 | 0.999 | -0.497 | 0.003 | 0.109 | 0.110 | 0.996 | -0.451 | 0.049 | 0.092 | 0.092 | 0.998 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

# Chapter 5

# Conclusion and Discussion

This paper has surveyed approaches to multicenter clinical trials for censored time to event data. When stratified analysis is required, we informally explain why a bias occurs if stratification is not considered. However, due to computational costs, it was not possible to determine the particular value of bias, but we did gain some insights into the direction of potential bias. According to the simulation results, the unadjusted Cox model was biased in all of the simulations. Therefore, multicenter clinical studies should use the stratified Cox or frailty models instead of the unadjusted Cox model.

We present the results of comparing the stratified Cox model and the gamma frailty model through simulation. We find that in a broad range of settings, the gamma frailty model, an example of the random effects approach, and the stratified Cox model produces estimates with lower bias than the unadjusted Cox model. It has been shown that, when survival data includes a some level of heterogeneity whatever its distribution, the two

model approach to estimate coefficients performs better than the unadjusted Cox proportional hazard model. The stratified Cox model does not generate bias, no matter which model is correct model. In the case of the frailty model, we found that there was no bias when the data was generated based on the gamma frailty model, and that bias existed when the data was generated on the stratified Cox model basis. It was also confirmed that the bias would large if a small number of samples were assigned in the stratum. However, the stratified Cox analysis showed that if the balance between treatment group and control group within the stratum was broken, the standard error would be larger than other models. In other words, stratified Cox model was able to confirm that it was well-fitting without significant impact on the form of baseline. In comparison, the frailty model was able to identify that the bias would occur if the multiply form on the baseline did not satisfied. Thus, model selection is important if many centers exists and the balance of the treatment group in the stratum is not satisfied.

We performed the additional simulations. For the simulation using data based on each model and strata were imbalanced, we assumed that the hazard would vary depending on the proportion of the imbalance. We supposed that center with many treatment groups have high or low levels of hazards. This is noted in the appendix. In fact, the situation is different for various centers. The larger the center, the more patients will be enrolled and the patient may experience better facility or nursing capacity. Therefore, the more patients in the center, the hazard may be low. Additional simulations are also needed that change the situation depending on the overall sample. Since the circumstances vary from center to center, this

should be considered when conducting actual multicenter clinical trials.

Further research is needed to determine which of these two models would be more appropriate for more complex situations. For example, in the presence of intra-cluster correlations (ICCs), even the random effects of individuals as well as the effects of centers could be considered. Also, in addition to the gamma frailty model, other frailty model could be considered. For example, the frailty model in which the frailty distribution is assumed to inverse Gaussian and positive stability density and so on.

# Appendix

Appendix table 1. Simulation using data based on the gamma frailty model when strata were imbalanced. (a)

(a) Center with many treatment group have high level of hazards.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **$\beta = -0.5$, Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.058 | 0.442 | 0.213 | 0.117 | 0.308 | -0.509 | -0.009 | 0.135 | 0.134 | 0.967 | -0.491 | 0.009 | 0.135 | 0.130 | 0.964 |
| | 8 | 0.028 | 0.528 | 0.172 | 0.117 | 0.187 | -0.497 | 0.003 | 0.134 | 0.136 | 0.964 | -0.467 | 0.033 | 0.133 | 0.131 | 0.947 |
| | 20 | 0.068 | 0.568 | 0.150 | 0.117 | 0.179 | -0.497 | 0.003 | 0.145 | 0.142 | 0.943 | -0.413 | 0.087 | 0.139 | 0.128 | 0.854 |
| | 100 | 0.102 | 0.602 | 0.121 | 0.117 | 0.152 | -0.503 | -0.003 | 0.168 | 0.171 | 0.853 | -0.184 | 0.316 | 0.144 | 0.126 | 0.275 |
| 2000 | 4 | -0.036 | 0.464 | 0.183 | 0.052 | 0.618 | -0.499 | 0.001 | 0.059 | 0.059 | 1.000 | -0.497 | 0.003 | 0.059 | 0.058 | 1.000 |
| | 8 | 0.020 | 0.520 | 0.139 | 0.052 | 0.475 | -0.502 | -0.002 | 0.058 | 0.059 | 1.000 | -0.499 | 0.001 | 0.057 | 0.059 | 1.000 |
| | 20 | 0.065 | 0.565 | 0.097 | 0.052 | 0.421 | -0.500 | 0.000 | 0.063 | 0.060 | 1.000 | -0.491 | 0.009 | 0.062 | 0.058 | 1.000 |
| | 100 | 0.095 | 0.595 | 0.064 | 0.052 | 0.451 | -0.497 | 0.003 | 0.065 | 0.063 | 1.000 | -0.422 | 0.078 | 0.062 | 0.057 | 1.000 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Appendix table 2. Simulation using data based on the gamma frailty model when strata were imbalanced. (b)

(b) Center with many treatment group have low level of hazards.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **$\beta = -0.5$, Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.786 | -0.286 | 0.147 | 0.123 | 1.000 | -0.506 | -0.006 | 0.137 | 0.134 | 0.967 | -0.522 | -0.022 | 0.136 | 0.131 | 0.980 |
| | 8 | -0.795 | -0.295 | 0.131 | 0.122 | 1.000 | -0.504 | -0.004 | 0.134 | 0.136 | 0.962 | -0.533 | -0.033 | 0.132 | 0.132 | 0.981 |
| | 20 | -0.813 | -0.313 | 0.136 | 0.123 | 1.000 | -0.508 | -0.008 | 0.150 | 0.142 | 0.935 | -0.579 | -0.079 | 0.142 | 0.132 | 0.988 |
| | 100 | -0.806 | -0.306 | 0.125 | 0.122 | 1.000 | -0.507 | -0.007 | 0.166 | 0.171 | 0.853 | -0.749 | -0.249 | 0.139 | 0.133 | 1.000 |
| 2000 | 4 | -0.780 | -0.280 | 0.092 | 0.054 | 1.000 | -0.502 | -0.002 | 0.058 | 0.059 | 1.000 | -0.504 | -0.004 | 0.057 | 0.058 | 1.000 |
| | 8 | -0.794 | -0.294 | 0.067 | 0.054 | 1.000 | -0.501 | -0.001 | 0.058 | 0.059 | 1.000 | -0.505 | -0.005 | 0.058 | 0.058 | 1.000 |
| | 20 | -0.804 | -0.304 | 0.063 | 0.054 | 1.000 | -0.502 | -0.002 | 0.060 | 0.060 | 1.000 | -0.512 | -0.012 | 0.058 | 0.058 | 1.000 |
| | 100 | -0.804 | -0.304 | 0.058 | 0.054 | 1.000 | -0.498 | 0.002 | 0.062 | 0.063 | 1.000 | -0.565 | -0.065 | 0.059 | 0.059 | 1.000 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Appendix table 3. Simulation using data based on the stratified Cox model when strata were imbalanced. (a)

(a) Center with many treatment group have high level of hazards.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| **$\beta = -0.5$, Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.257 | 0.243 | 0.120 | 0.117 | 0.583 | -0.500 | 0.000 | 0.134 | 0.134 | 0.969 | -0.384 | 0.116 | 0.131 | 0.123 | 0.859 |
| | 8 | -0.286 | 0.214 | 0.121 | 0.117 | 0.685 | -0.502 | -0.002 | 0.133 | 0.136 | 0.956 | -0.367 | 0.133 | 0.129 | 0.120 | 0.840 |
| | 20 | -0.296 | 0.204 | 0.116 | 0.117 | 0.726 | -0.498 | 0.002 | 0.142 | 0.142 | 0.932 | -0.343 | 0.157 | 0.122 | 0.118 | 0.822 |
| | 100 | -0.309 | 0.191 | 0.119 | 0.117 | 0.752 | -0.508 | -0.008 | 0.173 | 0.171 | 0.863 | -0.331 | 0.169 | 0.128 | 0.117 | 0.786 |
| 2000 | 4 | -0.256 | 0.244 | 0.052 | 0.052 | 0.998 | -0.500 | 0.000 | 0.059 | 0.059 | 1.000 | -0.434 | 0.066 | 0.053 | 0.057 | 1.000 |
| | 8 | -0.283 | 0.217 | 0.051 | 0.052 | 1.000 | -0.503 | -0.003 | 0.056 | 0.059 | 1.000 | -0.439 | 0.061 | 0.055 | 0.057 | 1.000 |
| | 20 | -0.296 | 0.204 | 0.052 | 0.052 | 1.000 | -0.502 | -0.002 | 0.059 | 0.060 | 1.000 | -0.410 | 0.090 | 0.058 | 0.055 | 1.000 |
| | 100 | -0.303 | 0.197 | 0.054 | 0.052 | 1.000 | -0.501 | -0.001 | 0.063 | 0.063 | 1.000 | -0.348 | 0.152 | 0.059 | 0.052 | 1.000 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

Appendix table 4. Simulation using data based on the stratified Cox model when strata were imbalanced. (b)

(b) Center with many treatment group have low level of hazards.

| N | # of strata | Unadjusted Cox | | | | | Stratified Cox | | | | | Gamma frailty | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power | Estimate | Bias | SD | SE | Power |
| $\boldsymbol{\beta = -0.5}$**, Censoring = 25%** | | | | | | | | | | | | | | | | |
| 400 | 4 | -0.599 | -0.099 | 0.121 | 0.123 | 0.999 | -0.495 | 0.005 | 0.133 | 0.134 | 0.958 | -0.476 | 0.024 | 0.131 | 0.129 | 0.958 |
| | 8 | -0.595 | -0.095 | 0.121 | 0.123 | 0.999 | -0.503 | -0.003 | 0.131 | 0.136 | 0.970 | -0.523 | -0.023 | 0.128 | 0.126 | 0.989 |
| | 20 | -0.591 | -0.091 | 0.126 | 0.123 | 0.998 | -0.503 | -0.003 | 0.140 | 0.142 | 0.944 | -0.563 | -0.063 | 0.130 | 0.124 | 0.991 |
| | 100 | -0.590 | -0.090 | 0.125 | 0.123 | 0.999 | -0.509 | -0.009 | 0.177 | 0.171 | 0.840 | -0.588 | -0.088 | 0.127 | 0.123 | 0.998 |
| 2000 | 4 | -0.602 | -0.102 | 0.057 | 0.055 | 1.000 | -0.499 | 0.001 | 0.058 | 0.059 | 1.000 | -0.438 | 0.062 | 0.056 | 0.059 | 1.000 |
| | 8 | -0.592 | -0.092 | 0.055 | 0.055 | 1.000 | -0.501 | -0.001 | 0.059 | 0.059 | 1.000 | -0.455 | 0.045 | 0.058 | 0.059 | 1.000 |
| | 20 | -0.591 | -0.091 | 0.054 | 0.055 | 1.000 | -0.504 | -0.004 | 0.059 | 0.060 | 1.000 | -0.493 | 0.007 | 0.057 | 0.057 | 1.000 |
| | 100 | -0.590 | -0.090 | 0.055 | 0.055 | 1.000 | -0.500 | 0.000 | 0.063 | 0.063 | 1.000 | -0.559 | -0.059 | 0.058 | 0.055 | 1.000 |

These values are mean of the estimates based on 1,000 replicates; SD: empirical standard deviation; SE: the estimated standard errors.

# References

Anderson, G. L., & Fleming, T. R. (1995). Model misspecification in proportional hazards regression. *Biometrika, 82*(3), 527-541.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *The annals of statistics*, 1100-1120.

Balan, T. A., & Putter, H. (2020). A tutorial on frailty models. *Statistical methods in medical research, 29*(11), 3424-3454.

Collett, D. (2014). *Modelling survival data in medical research* (3rd ed.). Chapman & Hall/CRC.

Colosimo, E. A. (2001). Stratification Points for the Proportional Hazards Model. *Brazilian Journal of Probability and Statistics*, 107-119.

David, G. K., & Mitchel, K. (2012). *Survival Analysis: A Self-Learning Text*. Spinger.

Glidden, D. V., & Vittinghoff, E. (2004). Modelling clustered survival data from multicentre clinical trials. *Statistics in medicine, 23*(3), 369-388.

Gutierrez, R. G. (2002). Parametric frailty and shared frailty survival models. *The Stata Journal, 2*(1), 22-44.

Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: techniques for censored and truncated data* (2th ed.). New York: Springer.

Localio, A. R., Berlin, J. A., Ten Have, T. R., & Kimmel, S. E. (2001). Adjustments for center in multicenter studies: an overview. *Annals of internal medicine, 135*(2), 112-123.

O'Quigley, J., & Stare, J. (2002). Proportional hazards models with frailties and random effects. *Statistics in medicine, 21*(21), 3219-3233.

Romdhane, S., & Belkacem, L. (2015). Frailty Modeling for clustered survival data: a simulation study. *Technical Report.*

Struthers, C. A., & Kalbfleisch, J. D. (1986). Misspecified proportional hazard models. *Biometrika, 73*(2), 363-369.

# 국 문 요 약

## 다기관 임상시험에서 군집화된 생존 자료에 대한 통계적 모델 비교

다국적 또는 다기관 임상시험을 수행하는 것은 약물 유효성을 테스트하는 인기 있는 방법이다. 기관의 수는 많지만 각 기관에는 소수의 환자만 포함되며, 하나의 기관에서 치료군과 대조군의 비율은 불균형적인 경우가 많다. 이러한 데이터를 분석할 때 통계 모델링을 통해 적절한 모델을 선택해야 한다. 또한, 다기관을 포함하는 연구는 층화를 고려한 분석을 수행해야 한다. 각 센터에 대해 기준 위험의 기능적 형태를 알 수 없는 경우, Glidden (2004)는 한계 위험 비율이 콕스 비례 위험 모델에서 편향되어 있음을 보여주었다. 다기관 임상시험에서 서로 다른 기관에서 다루는 환자들과 같이 환자의 특정 하위 그룹 사이에 차이가 있을 수 있다. 경우에 따라 부분군 효과의 영향에 대한 조정은 계획된 분석의 필수적인 부분이다. (ICH-E9) 따라서 각 기관에서 환자의 이러한 군집화된 데이터를 분석하려면 더 복잡한 방법이 필요하다.

이 상황에서는 층화된 콕스 모델 또는 프레일티 모델을 고려할 수 있다. 층화된 콕스 모델은 각 기관을 하나의 층으로 간주하면서 기관별로 층화한다. 프레일티 모형은 동일한 기관에 있는 환자가 다른 기관에 있는 환자보다 서로 더 유사하기 때문에 각 기관의 효과를 랜덤 효과로 나타낸다. 앞에서 언급한 두 모델은 다양한 시나리오에서 비교되었다. 다기관 설정에서 고려된 시나리오는 1) 총 표본 크기와 서로 다른 층수의 수, 2) 각 기관에서 치료군과 대조군의 비율, 3) 기관에서 기관까지 변화하는 기준 위험의 기능 형태, 그리고 4) 관측 중단 백분율이 달랐다. 각 시나리오에서 두 모델의 통계적 성능을 비교했다.

이 연구는 두 가지 주요 목표를 가지고 있다. 첫 번째 목적은 층화를 고려한 분석이 올바른 절차일 때 층화하지 않은 분석에서 추정치가 편향되는 이유를 확인하는 것이다. 두 번째는 다양한 시나리오에서 층화된 콕스 모델 또는 랜덤 효과 모델 중 어떤 모델이 더 적합한지 시뮬레이션을 통해 보여주는 것이다.

결론적으로, 우리는 시뮬레이션을 통해 통계 모델을 비교함으로써 어떤 모델이 어떤 맥락에서 더 적합할지 제시하려고 한다. 이 연구는 다국적 또는 다기관 임상시험을 수행할 때 고려할 수 있는 모델에 대한 전반적인 이해와 모델 선택에 도움이 될 것으로 기대된다.

---

핵심되는 말 : 생존 분석, 콕스 비례 위험 모델, 층화 콕스 모델, 감마 프레일티 모델, 다기관 임상시험