2016

Sample Size Calculation For PH Mixture Cure Model

Yihong Zhan
University of South Carolina

SAMPLE SIZE CALCULATION FOR PH MIXTURE CURE MODEL

by

Yihong Zhan

Bachelor of Science
Zhejiang University, 2004

Doctor of Philosophy
Purdue University, 2012

Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Science in Public Health in

Biostatistics

The Norman J. Arnold School of Public Health

University of South Carolina

2016

Accepted by:

Jiajia Zhang, Director of Thesis

Bo Cai, Reader

James W. Hardin, Reader

Lacy Ford, Senior Vice Provost and Dean of Graduate Studies

www.manaraa.com

ACKNOWLEDGEMENTS

Many people have helped make my time at USC a fantastic one. First, I would like to thank my advisor Dr. Jiajia Zhang, for her patient guidance, invaluable advice and constant inspiration. Her guidance on my thesis has allowed me to explore the idea of mixture cure models and to gain a deeper of understanding of sample size calculation knowledge. It has been my privilege to conduct this work under her direction.

I would like to extend my sincere gratitude to my committee members: Dr. James Hardin and Dr. Bo Cai for their insightful suggestions and generous encouragement. I truly want to thank both of you for the extraordinary amount of time, patience and knowledge you are willing to provide during my research.

I would also like to thank the faculty and staff of Department of Epidemiology and Biostatistics for sharing their knowledge and passion in Biostatistics with me. They're very supportive and accommodative.

Finally, I'm deeply grateful to my parents, my husband and my two lovely kids for their continuous love and support. I could not have done it without you.

ABSTRACT

With the development of advanced medical technology, a significant proportion of patients can be cured of many chronic diseases. Because a substantial fraction of patients have censored information, the standard survival model, such as the proportional hazards (PH) model cannot capture the cured information of patients. Thus PH mixture cure model is developed to handle the survival data with potential cured information. A corresponding sample size formula based on log rank test has been proposed by Wang et al. (2012) and the probability of death in their formula is only contributed by the control arm. However, to calculate the sample size and power, the hazard ratio and odds ratio are prespecified, which can also contribute to increase the accuracy of probability of death, by accounting for both the control and treatment arm. Therefore, we modify this formula by improving the estimation of the probability of death based on both the control and treatment arm using two approaches. The Schoenfeld and Ewell method adjusts the probability of death by averaging the death from both groups. The modified approach is verified by extensive simulation under exponential, weibull and lognormal distribution. The performance of the three methods has been compared under each setting with parametric and nonparametric estimation. Furthermore, the sample size calculation has been extended to PH mixture cure model with nonbinary covariates and evaluated by simulation studies. These modifications have been implemented in the R package and applied to the real data sets.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

In randomized controlled clinical trials with time-to-event endpoints, a sample size calculation is a very important consideration. Because the main goal in survival analysis is to distinguish the treatment effect among different groups, the number of patients enrolled should be adequate for obtaining the conclusion from statistical inference. On the other hand, excess number of patients may result in waste of resources. Thus developing accurate and widely used sample size calculation method has attracted more and more interests.

1.1 PROPORTIONAL HAZARDS MODEL

The proportional hazards (PH) model is one of the most popular models in survival analysis [1, 2]. The PH model can be expressed as

$$h(t|\boldsymbol{x}) = h_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{x})$$

where $h(\cdot)$ is the hazard function, $h_0(\cdot)$ is the baseline hazard function, $\boldsymbol{x}$ is a vector of covariates and $\boldsymbol{\beta}$ is an unknown vector of coefficients of interest. In the case where $\boldsymbol{x} = 0$, the hazard function $h(t|\boldsymbol{x})$ reduces to the baseline hazard function $h_0(t)$.

Under the PH model, the hazard ratio of two subjects is constant over time. For example, the hazard ratio of two subjects with covariates $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ is calculated as

$$HR = \frac{h_0(t)\exp(\beta x_1)}{h_0(t)\exp(\beta x_2)} = \exp[\boldsymbol{\beta}(\boldsymbol{x_1} - \boldsymbol{x_2})]$$

which is apparently constant over time. We could interpret the HR ($\exp[\boldsymbol{\beta}(\boldsymbol{x_1} - \boldsymbol{x_2})]$) as

1

the relative risk of an individual with risk factor x_1 having the event as compared to an individual with risk factor x_2 having the event. Particularly, given all other variables at the same level, when x is a binary indicator variable of control and treatment arm, $\exp(\boldsymbol{\beta})$ can be interpreted as the hazard ratio between two arms. Having this nice interpretation, the PH model is used for investigating the covariate effects on the hazard function.

The proportional hazards model has been widely used in practice due to its easy estimation through the nonparametric approach. Particularly, there is no need to specify the baseline hazard function and the MLE approach is approximated by the partial likelihood. The estimation based on this approach is consistent and satisfies the normal approximation. Based on the estimated parameter β, the survival function can be estimated nonparametrically. The partial likelihood has been implemented in most statistical software such as "phreg" in SAS, "coxph" in R, and "stcox" in stata [3],

1.2 SAMPLE SIZE CALCULATION FOR PROPORTIONAL HAZARDS MODEL

Especially for x being a binary variable, if letting x indicate the treatment effect ($x = 1$ for treatment arm and $x = 0$ for control arm), the PH model can be expressed as

$$h(t|x) = h_0(t)\exp(\beta_0 x)$$

After adjusting all the other covariates, $\exp(\beta_0)$ can be interpreted as the risk ratio of an individual in treatment arm having the event as compared to an individual in control arm having the event.

Assuming $1 - \theta$ is the power of statistical test, α is the significance level, and thus Z_θ and $Z_{\alpha/2}$ are the θ and $\alpha/2$ percentiles of the standard normal distribution, respectively; p is the proportion of subjects in the treatment arm; β_0 is the log-hazard ratio between

2

treatment and control arms; δ is the censoring indicator, for which 1 denotes failure and 0 denotes censoring. $P(\delta = 1)$ is the probability of failure. The following sample size formula based on the standard PH model has been widely used [4, 5],

$$n = \frac{(Z_\theta + Z_{\alpha/2})^2}{p(1-p)\beta_0^2 P(\delta=1)} \tag{1.1}$$

The major statistics software can perform sample size and power calculation for survival studies. In SAS [6], "TWOSAMPLESURVIVAL" statement in the POWER procedure can compare two survival curves based on log-rank test, using Schoenfeld's method [4]. Stata has several "stpower" commands for calculating sample size under various scenarios. The R package "powerSurvEpi" has several functions to calculate sample size for Cox PH model with both binary and nonbinary covariates [7]. However, these software packages have limited options in either the survival distribution function or the accrual pattern.

1.3 PROPORTIONAL HAZARDS MIXTURE CURE MODEL

As we know, in all the standard survival models, there's a common assumption, that all patients will eventually experience the event of interest, which means that when time is long enough, the survival probability for the population will become 0. However, with the development of advanced medical technology, many fatal and chronic diseases are now becoming curable. In this scenario, those cured patients will never experience the event of interest, even the follow-up period is sufficient. To account for this curable part, mixture cure model has been specifically designed to be applied for this type of study [8-10]. Different from the standard survival model, one advantage of the mixture cure model

3

is that it has two separate components to model the cure probability for the cured part and the survival probability for the uncured patients, respectively. The most popular mixture cure model for analyzing survival data was proposed by Boag [11], which could estimate the cure rate of the cured patients and the survival probability of uncured patients in the meantime.

Let T be a nonnegative random variable denoting the observed time of a patient, which is the minimum of the failure time and censoring time. We assume that T is independent of censoring. Let $S(t|x, z)$ denote the survival functions of T, where x and z are observed values of two covariate vectors on which the distribution of T may depend. The mixture cure model is then given by

$$S(t|x, z) = \pi(z) + (1 - \pi(z))S_u(t|x) \tag{1.2}$$

where $S(t|x, z)$ denotes the overall survival function. $\pi(z)$ is usually referred to as the "incidence" part, which is the probability of being cured giving predictors z via the inverse logistic function,

$$\pi(z) = \frac{\exp(\gamma' z)}{1 + \exp(\gamma' z)} \tag{1.3}$$

$S_u(t|x)$ is usually referred as "latency" part, which is the survival function of the uncured subjects depending on x. Under the proportional hazards model assumption, that is

$$S_u(t|x) = S_0(t|x)^{e^{\beta x}}$$

the above model (1.2) and (1.3) become PH mixture cure model.

There's an existing R package "*smcure*" developed for analyzing mixture cure models [12]. The package can estimate semiparametric PH mixture cure model and semiparametric accelerating failure time (AFT) mixture cure model, based on the EM

4

algorithm. In SAS, there's an existing macro called *PSPMCM* which accounts for the PHMC model and the parametric approach for the AFTMC model [13].

1.4 SAMPLE SIZE CALCULATION FOR PROPORTIONAL HAZARDS MIXTURE CURE MODEL

For mixture cure models, the most common case is that the covariate is binary variable, such as only including the control and treatment arm. If let $j = 0, 1$ denote the control and treatment arm; π_0, π_1 denote cure rate for control arm and treatment arm; $\lambda_0(t), \lambda_1(t)$ denote the hazard function for control arm and treatment arm; β_0 denote the log-hazard ratio between two arms and γ_0 denote the log-odds ratio between two arms. $S_j^*(\bullet)$ Denotes the overall survival function and $S_j(\bullet)$ denote the survival function of uncured patients. The mixture cure model is as follows,

$$S_j^*(t) = \pi_j + (1 - \pi_j)S_j(t)$$

$$\lambda_1(t) = e^{\beta_0}\lambda_0(t)$$

which accounts for the constant hazard ratio.

$$logit(\pi_1) = logit(\pi_0) + \gamma_0$$

which accounts for the constant odds ratio. In a special case that, when $\pi_0 = \pi_1 = 0$, it reduces to the standard PH model.

For a survival trial with a proportion of patients being cured, the hypothesis of interest is listed as follows,

$$H_0 : \beta_0 = \gamma_0 = 0 \quad \text{VS} \quad H_a : at\ least\ one\ of\ (\beta_0, \gamma_0) \neq 0$$

The alternative hypothesis H_a includes the following cases,

5

www.manaraa.com

1) $\beta_0 \neq 0$, $\gamma_0 \neq 0$: treatment has effects on both cure rate and survival probability of uncured patients;

2) $\beta_0 \neq 0$, $\gamma_0 = 0$: treatment only has effect on survival probability of uncured patients;

3) $\beta_0 = 0$, $\gamma_0 \neq 0$: treatment only has effect on cure rate.

Wang et al. developed a sample size formula for the PH mixture cure model based on the log rank test [14]:

For power $1 - \theta$ and significance level α,

$$n = \frac{(Z_\theta + Z_{\alpha/2})^2 \int_0^\infty S_C(t) f_0(t) dt}{p(1-p)(1-\pi_0)\left\{\int_0^\infty m(\gamma_0, \beta_0, \pi_0) S_C(t) f_0(t) dt\right\}^2} \tag{1.4}$$

where $m(\gamma_0, \beta_0, \pi_0) = \pi_0\{\gamma_0 + \Lambda_0(t)\beta_0\}/S_0^*(t) - \beta_0$

p is the proportion of subjects in the treatment arm; π_0 is the cure rate for control arm. β_0 denotes the log-hazard ratio between two arms and γ_0 denotes the log-odds ratio between two arms. $S_C(t)$ is the survival function of censoring time and $f_0(t)$ is the density function of survival times for uncured patients in the control arm. $S_0^*(t)$ is the overall survival function in the control arm.

When $\pi_0 = 0$, there's no cured part, and (1.4) reduces to the sample size formula for standard PH model in (1.1).

Therefore, this sample size formula extends the original formula for standard PH model by allowing a cure fraction. In Wang's paper [14], effects of the difference in the short-term survival and the difference in the cure fraction on the sample size has been studied, which are usually of main interest in many survival trials. Impacts of various

6

accrual rate patterns, including uniform, increasing, and decreasing, and also the length of accrual and follow-up periods on the estimated sample size have been illustrated.

Cai. et al implement the above sample size calculation method in R and develop a R package NPHMC [15]. In their package, they include many options as discussed in the paper. First, they provide option for accrual pattern, such as uniform, increasing, and decreasing; Secondly, the base hazard function for treatment is also flexible, which includes two parametric distributions, exponential and Weibull; Thirdly, in their output interface, either the sample size or power can be reported by given either power or sample size and power curves can be generated for multiple sample sizes for power analysis; Lastly, the package can either do the parametric estimation based on the known parameters, or do the nonparametric estimation from the available dataset.

1.5 MOTIVATION AND OUTLINE OF THESIS

Based on Wang's method, the sample size for the PH mixture cure model can be conveniently calculated. However, in the original method, the probability of death is contributed only from the control arm, which may either underestimate or overestimate the power. It's expected that when calculating the probability of death, if it can be accounted for in both arms, sample size calculation will be more precise comparing with the current method. The outline of the thesis is listed as follows,

In Chapter 2, we introduce the method of modification for sample size calculation under PH mixture cure model, using both the Schoenfeld [4] and Ewell [16] method. Secondly, the formula was updated to be adapted to cases with lognormal distribution in hazard function. Thirdly, formulas are developed to be used for nonbinary cases based on

7

Hsieh and Lavori's method [17], under exponential, Weibull and lognormal distribution in hazard function. All the methods are illustrated with notation specification, formulas and implementation in R.

In Chapter 3, extensive simulation are conducted to check the performance of the proposed methods. We first compare the performance of the three methods under different distribution in hazard function, by comparing the empirical power with the nominal power. Secondly, for each method, parametric estimation has been compared with nonparametric estimation.

In Chapter 4, the proposed method has been applied to the melanoma dataset from ECOG phase III clinical trial E1684. Furthermore, power analysis has been conducted using the proposed method, under various settings.

Finally in Chapter 5, we summarize all the work and present some discussion and conclusions for future work.

CHAPTER 2 METHODS AND SOFTWARE IMPLEMENTATION

In this chapter, methods used for formula modification and implementation are illustrated. Section 2.1 introduces the motivation and presents our two proposed modification methods; section 2.2 extends the current method and the two proposed method to lognormal distribution; section 2.3 demonstrates the components specification for how to apply all three methods in sample size calculation, under three distributions (exponential, Weibull and lognormal); section 2.4 develops the nonparametric estimation method using the two proposed methods; section 2.5 demonstrate the implementation of the three methods in R; section 2.5 further extends the methods to the sample size calculation for nonbinary cases.

2.1 SCHOENFELD AND EWELL METHOD MODIFICATION

For the sample size formula proposed by Wang in (1.4), the probability of death ($P\left[\delta = 1\right]$) is calculated only from the control arm, which may either underestimate or overestimate the necessary sample size.

For sample size calculation in standard PH model, Schoenfeld [4] proposed a method to adjust the calculation of the probability of death, based on both the control and the treatment arm, with p denoting the proportion of subjects in the treatment arm.

$$P(\delta = 1) = pP(\delta = 1|Treatment) + (1 - p)P(\delta = 1|Control)$$

We refer to this method as Schoenfeld method. If we expand the formula, it is as follows,

$$n = \frac{(Z_\theta + Z_{\alpha/2})^2}{p(1-p)\beta_0^2\{p\int_0^\infty S_C(t)f_1(t)\,dt + (1-p)\int_0^\infty S_C(t)f_0(t)\,dt\}}$$

It is apparent that it accounts for the different density function between the control and treatment arm, which is closer to the real situation.

Instead of $f_0(t)$ in Wang's method, $\bar{f}(t)$ is applied to formula which is an average probability density function from both the control arm and treatment arm.

$$\bar{f}(t) = pf_1(t) + (1-p)\,f_0(t)$$

By substituting the f_0 in Wang's method by $\bar{f}$, it becomes the modified sample size calculation formula, which is written as

$$n = \frac{(Z_\theta + Z_{\alpha/2})^2 \int_0^\infty S_C(t)\bar{f}(t)dt}{p(1-p)(1-\pi_0)\left\{\int_0^\infty \bar{m}(\gamma_0, \beta_0, \pi_0)S_C(t)\bar{f}(t)dt\right\}^2} \tag{2.1}$$

where $\bar{m}(\gamma_0, \beta_0, \pi_0) = \pi_0\{\gamma_0 + \bar{\Lambda}(t)\beta_0\}/\bar{S}^*(t) - \beta_0$.

As for the numeric average of density function, other components can be derived similarly, which are listed as follows,

- $\bar{f}(t) = pf_1(t) + (1-p)f_0(t),$ $\bar{S}(t) = pS_1(t) + (1-p)S_0(t)$

- $\bar{S}^*(t) = \pi_0 + (1-\pi_0)\bar{S}_0(t),$ $\bar{\Lambda}(t) = -log\bar{S}(t)$

Besides the Schoenfeld method which averages the density function first, an alternative method proposed by Ewell [16] which aims to average the hazard function first and we refer to it as Ewell method. First, the weighted hazard function was calculated from the square root of the product of hazard function from both control and treatment arms. Then the density function and survival function is calculated based on the hazard function. We listed each component as follows:

Ewell method:

$\Lambda_i(t)$: Cumulative hazard function of the ith group, where $i = 0,1$

- $\bar{\Lambda}(t) = \Lambda_1(t)^p \Lambda_0(t)^{1-p}, \qquad\qquad \bar{S}(t) = \exp\{-\bar{\Lambda}(t)\}$

- $\bar{S}^*(t) = \pi_0 + (1 - \pi_0)\bar{S}_0(t), \qquad \bar{f}(t) = -\frac{\partial \bar{S}(t)}{\partial t}$

Similar as Schoenfeld method, f_0 in Wang's method is substituted by $\bar{f}$, which is calculated based on Ewell method. Then the other components calculated as above are substituted into the formula.

2.2 LOGNORMAL DISTRIBUTION

As mentioned before, the existing method and package only cover exponential and Weibull distribution, in our study all three methods has been accommodated to lognormal distribution.

Formulas for lognormal distribution in R will be utilized here. Similar as for the exponential and Weibull distributions, the Wang method calculates the probability of death based on the control arm, while Schoenfeld method acquires the weighted functions from the numeric average of both the control and treatment arm. The Ewell method uses the geometric average to get necessary functions. Some related formulas are as follows,

➢ Wang method:

$$S_0(t) = 1 - pnorm(\frac{\ln t - \mu_0}{\sigma_0}), \qquad S_0{}^*(t) = \pi_0 + (1 - \pi_0)S_0(t),$$

$$f_0(t) = dnorm\left(\frac{\ln t - \mu_0}{\sigma_0}\right)\frac{1}{\sigma_0}\frac{1}{t}, \qquad \Lambda_0(t) = -\log(S_0(t)).$$

➢ Schoenfeld method:

$$S_0(t) = 1 - pnorm(\frac{\ln t - \mu_0}{\sigma_0}), \qquad S_1(t) = S_0(t)^{e^{\beta}},$$

11

$$f_0(t) = dnorm\left(\frac{\ln t - \mu_0}{\sigma_0}\right)\frac{1}{\sigma_0}\frac{1}{t}, \qquad f_1(x) = e^\beta S_0(t)^{e^\beta - 1} f_0(t)$$

> ➢ Ewell method:

$$S_0(t) = 1 - pnorm(\frac{\ln t - \mu_0}{\sigma_0}), \qquad S_1(t) = S_0(t)^{e^\beta},$$

$$\Lambda_0(t) = -\log(S_0(t)), \qquad \Lambda_1(t) = -\log(S_1(t)),$$

$$\bar{\Lambda}(t) = \Lambda_1(t)^p \Lambda_0(t)^{1-p}, \qquad \bar{S}(t) = \exp(-\bar{\Lambda}(t)),$$

$$\bar{S}^*(t) = \pi_0 + (1 - \pi_0)\bar{S}(t), \qquad \bar{f}(t) = \exp(-\bar{\Lambda}(t))\frac{\partial \bar{\Lambda}(t)}{\partial t}.$$

2.3 COMPONENT SPECIFICATIONS

To calculate the sample size using the formula shown above, several components need to be specified, such as survival function of censoring time, corresponding to different accrual patterns. There are three patterns considered here, uniform, increasing and decreasing. Specifically, if we let τ_a and τ_f denote the accrual and the follow-up period, respectively; $g(t)$ denote the probability density function of accrual times. Their corresponding survival function of the censoring times are summarized in the table below, which will be substituted into the formula.

For the density function of survival probability, we'll consider three different parametric distributions, including exponential, Weibull and lognormal. The lognormal distribution is the new feature we add to the original package. Some expressions for the three distributions are as follows:

Exponential Distribution

$$\lambda_0(t) = \lambda_0 k, \qquad \Lambda_0(t) = \lambda_0 t, \qquad S_0(t) = e^{-\lambda_0 t},$$

$$f_0(t) = \lambda_0 \exp(-\lambda_0 t), \quad for\ t > 0.$$

12

Table 2.1 Density Functions g(t) of Accrual Times and Survival Functions $S_c(t)$ of Censoring Times

Accrual	$g(t)$	$S_c(t)$
Uniform	$g(t) = \begin{cases} \frac{1}{\tau_a}, & if\ 0 < t \le \tau_a \\ 0, & otherwise \end{cases}$	$S_c(t) = \begin{cases} 1, & if\ t \le \tau_f \\ \frac{\tau_a + \tau_f - t}{\tau_a}, & if\ \tau_f < t \le \tau_a + \tau_f \\ 0, & if\ t > \tau_a + \tau_f \end{cases}$
Increasing	$g(t) = \begin{cases} \frac{2t}{\tau_a^2}, & if\ 0 < t \le \tau_a \\ 0, & otherwise \end{cases}$	$S_c(t) = \begin{cases} 1, & if\ t \le \tau_f \\ \frac{(\tau_a + \tau_f - t)^2}{\tau_a^2}, & if\ \tau_f < t \le \tau_a + \tau_f \\ 0, & if\ t > \tau_a + \tau_f \end{cases}$
Decreasing	$g(t) = \begin{cases} \frac{2(\tau_a - t)}{\tau_a^2}, & if\ 0 < t \le \tau_a \\ 0, & otherwise \end{cases}$	$S_c(t) = \begin{cases} 1, & if\ t \le \tau_f \\ 1 - \frac{(\tau_f - t)^2}{\tau_a^2}, & if\ \tau_f < t \le \tau_a + \tau_f \\ 0, & if\ t > \tau_a + \tau_f \end{cases}$

Weibull Distribution

$$\lambda_0(t) = \lambda_0 k (\lambda_0 t)^{k-1}, \qquad \Lambda_0(t) = (\lambda_0 t)^k, \qquad S_0(t) = e^{-(\lambda_0 t)^k},$$

$$f_0(t) = \lambda_0 k (\lambda_0 t)^{k-1} \exp(-(\lambda_0 t)^k), \quad for\ t > 0.$$

Lognormal Distribution

$$S_0(t) = 1 - pnorm(\frac{lnt - \mu_0}{\sigma_0}), \quad f_0(t) = dnorm\left(\frac{lnt - \mu_0}{\sigma_0}\right) \frac{1}{\sigma_0} \frac{1}{t}$$

To apply the Schoenfeld or Ewell method is to calculate $\bar{f}(t), \bar{S}(t), \bar{S}^*(t)$ and $\bar{\Lambda}(t)$ based on the above equations, discussed in section 2.2, respectively. Then those calculated average functions are substituted into the formula (2.1). For example, if a trial has uniform distribution for the accrual period, for power $1 - \theta$ and significance level α, the sample size can be calculated as the following,

$$n = \frac{(z_\theta + z_{\alpha/2})^2 \left\{ \int_0^{\tau_f} \bar{f}(t)dt + \int_{\tau_f}^{\tau_a + \tau_f} \frac{\tau_a + \tau_f - t}{\tau_a} \bar{f}(t)dt \right\}}{p(1-p)(1-\pi_0) \left\{ \int_0^{\tau_f} \bar{m}(\gamma_0, \beta_0, \pi_0)\bar{f}(t)dt + \int_{\tau_f}^{\tau_a + \tau_f} \frac{\tau_a + \tau_f - t}{\tau_a} \bar{m}(\gamma_0, \beta_0, \pi_0)\bar{f}(t)dt \right\}^2}$$

where $\bar{m}(\gamma_0, \beta_0, \pi_0) = \pi_0 \{\gamma_0 + \bar{\Lambda}(t)\beta_0\}/\bar{S}^*(t) - \beta_0$.

2.4 EXAMPLES FOR NONPARAMETRIC ASSUMPTION

Alternative to parametric assumptions, the sample size for nonparametric assumptions could also be calculated. If a dataset is available, the baseline survival function, cure rate and hazard ratio could be first estimated by an existing R package, *smcure*, which is developed for mixture cure model. By specifying the significance level, power and accrual pattern, the sample size can be attained as follows,

$$n = \frac{\left(z_\theta + z_{\alpha/2}\right)^2 \sum_{i=1}^{k} \hat{S}_0\left(t_{(i)}\right) S_c(t_{(i)})}{p(1-p)(1-\pi_0)\left\{\sum_{i=1}^{k} \hat{S}_0\left(t_{(i)}\right) S_c(t_{(i)}) \hat{m}(\hat{\gamma}_0, \hat{\beta}_0, \hat{\pi}_0; t_{(i)})\right\}^2}$$

where $\hat{m}(\hat{\gamma}_0, \hat{\beta}_0, \hat{\pi}_0; t_{(i)}) = \hat{\pi}_0\{\hat{\gamma}_0 + \hat{\Lambda}(t)\hat{\beta}_0\}/\hat{S}^*(t) - \hat{\beta}_0$

$$\hat{\Lambda}_0(t) = -log\hat{S}_0(t), \qquad \hat{S}_0^*(t) = \hat{\pi}_0 + (1-\hat{\pi}_0)\hat{S}_0(t)$$

Similar to the parametric one, numeric or geometric average will be calculated for Schoenfeld and Ewell method as following,

Schoenfeld method:

$\hat{S}_0(t)$: estimated from *smcure* package,

$$\hat{S}_1(t) = \hat{S}_0(t)^{e^{\hat{\beta}_0}}, \qquad \hat{S}(t) = p\hat{S}_1(t) + (1-p)\hat{S}_0(t),$$

$$\hat{S}_0^*(t) = \hat{\pi}_0 + (1-\hat{\pi}_0)\hat{S}_0(t), \qquad \hat{\Lambda}(t) = -\log(\hat{S}(t)).$$

Ewell method:

$\hat{S}_0(t)$: estimated from *smcure* package,

$$\hat{\Lambda}_0(t) = -\log(\hat{S}_0(t)), \qquad \hat{S}_1(t) = \hat{S}_0(t)^{e^{\hat{\beta}_0}},$$

$$\hat{\Lambda}_1(t) = -\log\left(\hat{S}_1(t)\right) = -e^{\hat{\beta}_0}\log(\hat{S}_0(t)), \qquad \hat{\Lambda}(t) = \hat{\Lambda}_1(t)^p \hat{\Lambda}_0(t)^{1-p},$$

$$\hat{S}(t) = \exp(-\hat{\Lambda}(t)), \qquad \hat{S}^*(t) = \hat{\pi}_0 + (1-\hat{\pi}_0)\hat{S}(t)$$

14

2.5 IMPLEMENTATION IN R

The sample size formula under the exponential, Weibull and lognormal distribution and also for the nonparametric estimation with different accrual patterns are implemented in "*allNPHMC*" package we developed. The function can be called using the following syntax:

allNPHMC <- function (n, power, alpha, accrualtime, followuptime, p, accrualdist= c("uniform", "increasing", "decreasing"), hazardratio, oddsratio, pi0, survdist = c("exp","weib","lgnorm"), shape, scale, data=NULL, output=NULL)

The arguments are as follows:

- n: the sample size needed for the power calculation.
- power: the power needed for sample size calculation, default is 90%.
- alpha: the level of significance of the statistical test, default is 0.05.
- accrualtime: the length of accrual period.
- followuptime: the length of follow-up time.
- p: the proportion of subjects in the treatment arm, default is 0.5.
- accrualdist: the accrual pattern. It can be "uniform","increasing" or "decreasing".
- hazardratio: the hazard ratio of uncured patients between two arms, which is defined as $e^{\beta_0} = \lambda_1(t)/\lambda_0(t)$. The value must be greater than 0 and not equal to 1.
- oddsratio: the odds ratio of cure rates between two arms, which is equivalent to $e^{\gamma_0} = [\pi_1/(1 - \pi_1)]/[\pi_0/(1 - \pi_0)]$. The value should be greater than 0 if there's cured fraction. When it is 0, the model is reduced to the standard proportional hazards model, which means there is no cure rate.
- pi0: the cure rate for the control arm, which is between 0 and 1.
- survdist: the survival distribution of uncured patients. It can be "exp", "weib" or "lgnorm".
- shape: shape parameter for distribution of survival function. By default shape=1, which refers to the exponential distribution.

15

- scale: scale parameter for distribution of survival function.
- data: if observed/historical data is available, the sample size can be calculated based on the nonparametric estimators from the PH mixture cure model by *smcure* package in R. The data must contain three columns with the order of "Time", "Status" and "X" where "Time" refers to observed time, "Status" refers to censoring indicator (1 = event of interest happens, and 0 = censoring) and "X" refers to arm indicator (1 = treatment and 0 = control). By default, data = NULL.
- method: It can be "wang","schoenfeld","ewell". if output is NULL, only result from Schoenfeld method will be output. To obtain the results from all three method (Wang, Schoenfeld, Ewell), method should be set to "all". By default, output = NULL.

Remarks:

"oddsratio = 1" can be specified if we believe the difference between two arms does not exist in the cure fraction. "hazardratio = 1" ($\beta_0 = 0$) can also be specified if we believe the difference between two arms does not exist in the uncured fraction. If the argument "data" is not "NULL", the "hazardratio" and the "oddsratio" will be automatically calculated based on the output from *smcure* package. If the argument "data" is not "NULL" and the "hazardratio" and the "oddsratio" are given, it will give a warning message as "The "hazardratio" and the "oddsratio" are not needed when the data option is specified. If the argument "data" is "NULL", the value of the "hazardratio" and the "oddsratio" need to be specified.

If power (sample size) and significant level of statistical test is given, the output of sample size (power) calculation is as follows:

If data = NULL, the output will display:

- PH Mixture Cure Model: n (Power)

16

- Standard PH Model: n (Power)

If data is specified: the output will display:

- Estimators from *smcure* package

- PH Mixture Cure Model: n (Power)

- Standard PH Model: n (Power)

2.6 SAMPLE SIZE CALCULATION FOR NONBINARY COVARIATE

In many experiments, the covariates are not binary variables, such as the covariates recording the dose levels, or the covariates being as risk factors, such as numbers of cigarettes smoked per day. To apply our sample size calculation method to this type of study, the above formula has been modified to accommodate to the nonbinary covariates. In 2000, Hsieh et al. proposed a sample size formula for nonbinary covariate without assuming any distribution on survival time. As seen from the above formula for standard PH model, the variance from the binary covariate contributes as $p(1-p)$. When applied to the nonbinary covariate, this variance item should be replaced by the variance of the nonbinary covariate, which is typically denoted as σ^2.

Thus, the sample size for standard PH model is calculated as the following,

$$n = \frac{(Z_\theta + Z_{\alpha/2})^2}{\sigma^2 \beta_0^2 P(\delta = 1)}$$

Based on above Hsieh's method, sample size for mixture cure model with nonbinary covariate could be calculated as follows,

➢ Wang method:

17

$$n = \frac{(Z_\theta + Z_{\alpha/2})^2 \int_0^\infty S_C(t) f_0(t) dt}{\sigma_0^2 (1-\pi_0) \left\{ \int_0^\infty m(\gamma_0, \beta_0, \pi_0) S_C(t) f_0(t) dt \right\}^2} \tag{15}$$

➤ Ewell method: (16)

Exponential and Weibull distribution:

$$\Lambda_0(t) = (\lambda_0 t)^k, \qquad \bar{\Lambda}(t) = \Lambda_0(t) e^{\beta \bar{x}}, \qquad \bar{S}(t) = \exp\left(-(\lambda_0 t)^k e^{\beta \bar{x}}\right)$$

$$\bar{f}(t) = -\frac{\partial \bar{S}(t)}{\partial t} = \bar{S}(t) e^{\beta \bar{x}} k (\lambda_0 t)^{k-1} \lambda_0$$

Lognormal distribution:

$$S_0(t) = 1 - pnorm\left(\frac{lnt - \mu_0}{\sigma_0}\right), \qquad \Lambda_0(t) = -\log(S_0(t)),$$

$$\bar{\Lambda}(t) = \Lambda_0(t) e^{\beta \bar{x}}, \qquad \bar{S}(t) = \exp[-\bar{\Lambda}(t)] = \exp[-\Lambda_0(t) e^{\beta \bar{x}}]$$

$$\bar{f}(t) = -\frac{\partial \bar{S}(t)}{\partial t} = \exp[-\bar{\Lambda}(t)] \, e^{\beta \bar{x}} \frac{1}{S_0(t)} dnorm\left(\frac{lnt - \mu_0}{\sigma_0}\right) \frac{1}{t} \frac{1}{\lambda_0}$$

Here the interpretation for the hazard ratio and odds ratio has changed. Compared to the binary case, both the hazard ratio and the odds ratio are now interpreted as change of the hazard function being associated with one unit change in corresponding covariate.

Comment: we only use the Ewell method, because its expression can be expanded to the continuous variable conveniently. For Schoenfeld method, to calculate the mean of the continuous case, it will involve the integration without the closed form. The numerical integration may be needed and thus currently we choose the easily implementable approach.

18

CHAPTER 3 SIMULATION STUDY

The purpose of this chapter is to conduct comprehensive simulation studies to investigate the performance of the proposed sample size calculations.

In section 3.1, we first evaluate the consistency of the nominal power and empirical power based on all three different methods (Wang, Schoenfeld and Ewell methods) under different simulation settings. The empirical power attained based on 2000 simulations will be compared to the nominal power which is set at 0.9 here. Performance will be evaluated based on the relative difference between the empirical power and the nominal power.

In section 3.2, performance of three methods will be evaluated by checking the consistency between parametric estimation and nonparametric estimation for each method. For each set of parameters, we compare the sample size from the fully specified distribution with from the fully nonparametric approach.

Lastly, in section 3.3, we extend the simulation to the continuous variable based on the formula provided in section 2.6. Wang and Ewell method has been applied. Simulation has been conducted to compare the performance of these two methods, under different distributions of hazard function, including exponential, Weibull and lognormal.

3.1 Simulation Comparing Wang, Schoenfeld and Ewell methods

3.1.1 Simulation Settings

EXPONENTIAL DISTRIBUTION

The following settings are used in the simulation study for exponential distribution. Covariate data are generated from a binomial distribution, which corresponding to control arm and treatment arm, with equal allocation.

We specifically consider the control and treatment arm follow the mixture cure model with the exponential distribution, which can be written as

$$S_0(t) = \pi_0 + (1 - \pi_0)e^{-\lambda_0 t}$$

$$S_1(t) = \pi_1 + (1 - \pi_1)e^{-\lambda_1 t}$$

Which leads to the odds ratio as $\log(\frac{\pi_0}{1-\pi_0}/\frac{\pi_1}{1-\pi_1})$ and hazard ratio as $\frac{\lambda_1}{\lambda_0}$.

1) The scale parameter is set at $\lambda_0 = 0.1, 0.2\ 0.3.$, which correspondes to the mean survival time of control group as 10, 5, 2.

2) For each λ_0, cure rate for control arm π_0 is set at 0.4 and 0.3.

The choice of $\lambda_1\ and\ \pi_1$ are determined by the odds ratio and hazard ratio, which is set up as: $e^{-\beta_0}$ is at (1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8), e^{γ_0} is at (0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0), which correspond to $\pi_1 = (0.211, 0.25, 0.286, 0.318,\ 0.348, 0.375, 0.4)$ for $\pi_0 = 0.4$ and $\pi_1 = (0.146, 0.176, 0.205, 0.231, 0.255, 0.278, 0.3)$ for $\pi_0 = 0.3$. For $\lambda_0 = 0.1$, λ_1 is varied at (0.083, 0.077, 0.071, 0.067, 0.063, 0.059, 0.056); For $\lambda_0 = 0.2$, λ_1 is varied at (0.167, 0.154, 0.143, 0.133, 0.125, 0.118, 0.111); For $\lambda_0 = 0.5$, λ_1 is varied at (0.417, 0.385, 0.357, 0.333, 0.313, 0.294, 0.278). And thus the mean survival time of treatment

arm is varied at (12, 13, 14, 15, 16, 17, 18), (6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0), (2.4, 2.6, 2.8, 3.0, 3.2, 3.4, 3.6).

To calculate the empirical power, the sample size at power 90% should be specified. The formula to calculate the sample size is reduced to

$$n = \frac{42.03}{\beta_0^2 P(\delta = 1)}$$

The censoring distribution is calculated based on section 2.3. For example, assuming uniform distribution in the accrual period with an accrual period of 1 year and a follow up time of 4 years; For $\pi_0 = 0.4$, the mean censoring rate is 0.83, 0.72, 0.56 for $\lambda_0 = 0.1, 0.2, 0.5$, respectively. For $\pi_0 = 0.3$, the mean censoring rate is 0.80, 0.67 and 0.48.

Weibull distribution

Settings used in the simulation study for Weibull distribution are listed as follows: Control and treatment group follow the mixture cure model with the Weibull distribution, which can be written as

$$S_0(t) = \pi_0 + (1 - \pi_0)e^{-(\lambda_0 t)^k}$$

$$S_1(t) = \pi_1 + (1 - \pi_1)e^{-(\lambda_1 t)^k}$$

Which leads to the odds ratio as $\log\left(\frac{\pi_0}{1-\pi_0}/\frac{\pi_1}{1-\pi_1}\right)$ and hazard ratio as $(\frac{\lambda_1}{\lambda_0})^k$.

1) The shape parameter k is set at 2. Scale parameter is set at $\lambda_0 = 0.1, 0.2\ 0.3$., which corresponding to the mean survival time of control group as 8.862, 4.431, 2.954.

2) For each λ_0, cure rate for control arm π_0 is set at 0.4 and 0.3.

The choice of $\lambda_1\ and\ \pi_1$ are determined by the odds ratio and hazard ratio, which are set up as: $e^{-\beta_0}$ is at (1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8), e^{γ_0} is at (0.4, 0.5, 0.6, 0.7, 0.8,

21

0.9, 1.0), which corresponding to $\pi_1 = (0.211, 0.25, 0.286, 0.318, \ 0.348, 0.375, 0.4)$ for

$\pi_0 = 0.4$ and $\pi_1 = (0.146, 0.176, 0.205, 0.231, 0.255, 0.278, 0.3)$ for $\pi_0 = 0.3$. For

$\lambda_0 = 0.1$, λ_1 is varied at $(0.091, 0.088, 0.085, 0.082, 0.079, 0.077, 0.075)$; For $\lambda_0 = 0.2$,

λ_1 is varied at $(0.183, 0.175, 0.169, 0.163, 0.158, 0.153, 0.149)$; For $\lambda_0 = 0.3$, λ_1 is varied

at $(0.274, 0.263, 0.254, 0.245, 0.237, 0.230, 0.224)$. And thus the mean survival time of

treatment group is varied at (9.708, 10.105, 10.486, 10.854, 11.210, 11.555, 11.890), (4.854,

5.052, 5.243, 5.427, 5.605, 5.777, 5.945), (3.236, 3.368, 3.495, 3.618, 3.737, 3.852, 3.963).

For $\pi_0 = 0.4$, the mean censoring rate is 0.92, 0.74, 0.59 for $\lambda_0 = 0.1, \ 0.2, \ 0.3$,

respectively. For $\pi_0 = 0.3$, the mean censoring rate is 0.90, 0.69 and 0.52.

Lognormal distribution

Besides Exponential and Weibull distribution, simulation is also conducted on

lognormal distribution. The specific settings are listed below:

Control and treatment arm follow the mixture cure model with the lognormal distribution,

which can be written as

$$S_0(t) = \pi_0 + (1 - \pi_0)(1 - pnorm\left(\frac{lnt - \mu_0}{\sigma_0}\right))$$

$$S_1(t) = \pi_1 + (1 - \pi_1)S_0(t)^{e^{\beta_0}}$$

Which leads to the odds ratio as $\log\left(\frac{\pi_0}{1-\pi_0} \middle/ \frac{\pi_1}{1-\pi_1}\right)$.

1) Scale parameter is set at $\sigma_0 = 1.5, 2.5, 3.5.$, which corresponding to the mean

    survival time of control group as 8.373, 61.868, 1242.648.

2) For each σ_0, cure rate for control arm π_0 is set at 0.4 and 0.3.

22

3) The setting of hazard ratio between two groups is the same as exponential and Weibull distribution. The setting of cure rates for control group and treatment group is the same as exponential and Weibull distribution.

4) The censoring distribution is calculated based on section 2.3. For example, assuming uniform distribution in the accrual period with an accrual period of 1 year and a follow up time of 15 years; For $\pi_0 = 0.4$, the mean censoring rate is 0.57, 0.64, 0.68 for $\sigma_0 = 1.5, 2.5, 3.5$, respectively. For $\pi_0 = 0.3$, the mean censoring rate is 0.49, 0.57 and 0.61.

3.1.2 SIMULATION RESULTS

For each set of data, censoring rate, estimated sample size, empirical power is reported. To evaluate the performance of all three methods, the absolute relative bias is reported, which is calculated as

$$\frac{|\text{empirical power} - 0.9|}{0.9} \times 100\%,$$

The absolute relative bias can measure the distance of the estimated power to the nominal power. A smaller relative bias indicates that the empirical value is closer to the nominal value, which is set at 0.9 here. For comparison purposes, the absolute relative biases from all methods are illustrated in Figure 3.1-3.6. In all the plots, the y-axis denotes the absolute relative bias, and x-axis denotes the hazard ratio. Three different methods are denoted by dotted lines, where Wang method is denoted as "Wang" in blue dotted line, Schoenfeld method is denoted as "Schoenfeld" in yellow dotted line and Ewell method is denoted as "Ewell" in red dotted line. The other specific settings are denoted by the label in each figure, which include the scale and shape parameters, accrual and follow up time, cure rate and

23

mean censoring rate. The three plots in the left panel is all for $\pi_0 = 0.4$, while the other three plots in the right panel is for $\pi_0 = 0.3$.

Figure 3.1 compares the three methods under exponential distribution. As for the three graphs in the left panel, with the same cure rate 0.4, censoring rate decreases from 0.83 to 0.56. With the decreasing censoring rate, the absolute relative bias (ARS) from Wang method experienced the large decrease. This trend can't be observed for the other two methods. Comparing all three methods, in relatively high censoring rate (censr = 0.83, left upper panel), the empirical power calculated by Wang's method largely under estimate the power. Compared with the Wang method, Schoenfeld and Ewell methods both work well, in which the largest deviation from the nominal power 0.9 is only 2%. There're not significantly differences between Schoenfeld and Ewell methods. When $\lambda_0 = 0.2$, censr = 072 (left, middle panel), Wang method has better performance than in the higher censoring rate. While for the other two methods, they both maintain good performance in all the estimation, deviation from the nominal power 0.9 are all smaller than 3%. In the scenario of even lower censoring rate (censr = 0.56, left low panel). There are not significant differences among all these three methods, deviation from the nominal power 0.9 are all smaller than 4%. Similar trends could be observed from the other three graphs in the right panel, in which $\pi_0 = 0.3$. From Figure 3.1, we conclude that both Schoenfeld and Ewell methods have improvements over the Wang method in the scenarios of high and medium censoring rates, in which they get more accurate sample size estimation. While for the low censoring rate case, all these three methods have comparable performance.

If there's longer follow up period, more events could be detected, which will certainly lower the censoring rate. Figure 3.2 investigates the performance of the three

24

methods under various following up period. The settings are as follows, $\pi_0 = 0.4, \lambda_0 = 0.1, \tau_a = 1$. τ_f is varied at (4, 10, 15, 20). The same sets of β_0/γ_0 are used as above. By
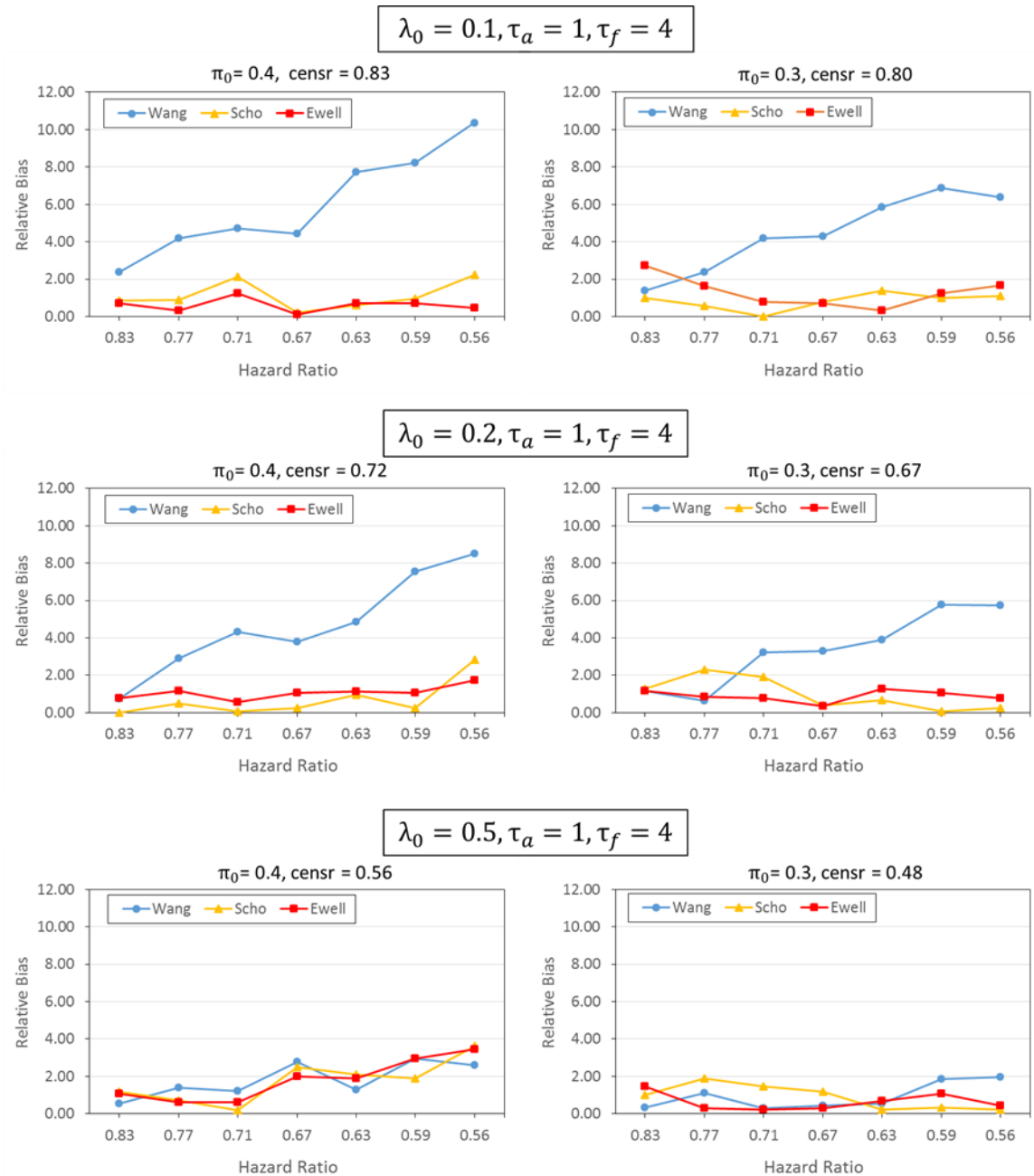
Figure 3.1 Simulation for Exponential distribution under parametric condition

adjusting τ_f, mean censoring rates vary from 0.5 to 0.8. The similar conclusion is achieved

as in Figure 3.1. Scho and Ewell methods could get better sample size estimation than

25

Wang method in the case of high censoring rate; while at low censoring rate, it couldn't be distinguished for the performance among these three methods. The differences between the empirical power and the nominal power are all less than 4%.
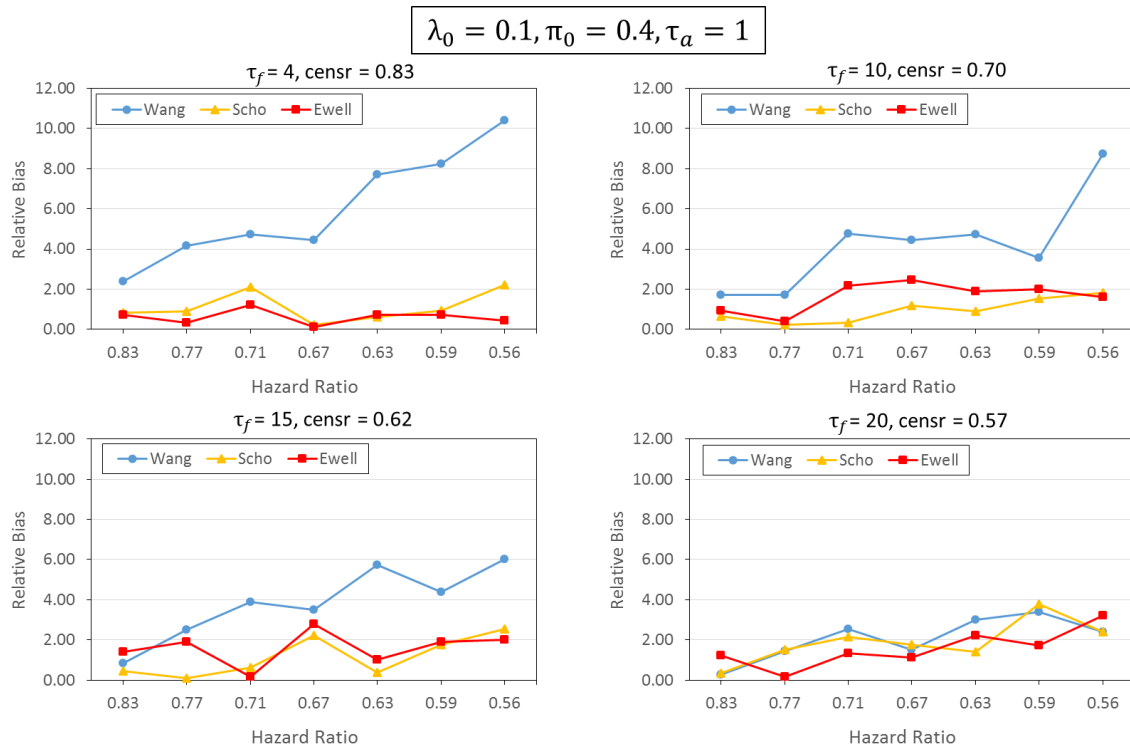
Figure 3.2 Simulation for Exponential distribution, adjusting follow up period

In Weibull distribution, increasing λ_0 results in shorter mean survival time and more event could be detected, and thus lower censoring rate could be acquired. The layout of Figure 3.3 is similar as Figure 3.1. As is shown here, both Schoenfeld and Ewell methods show superior performance over Wang method in all scenarios. Most of the results form Schoenfeld and Ewelll methods show less than 2% deviation from the nominal power 0.9, while the results from Wang method show as high as 10% deviation from 0.9. The

26

differences between Wang and the other two methods are larger in case of higher censoring

rate than in the lower censoring rate.
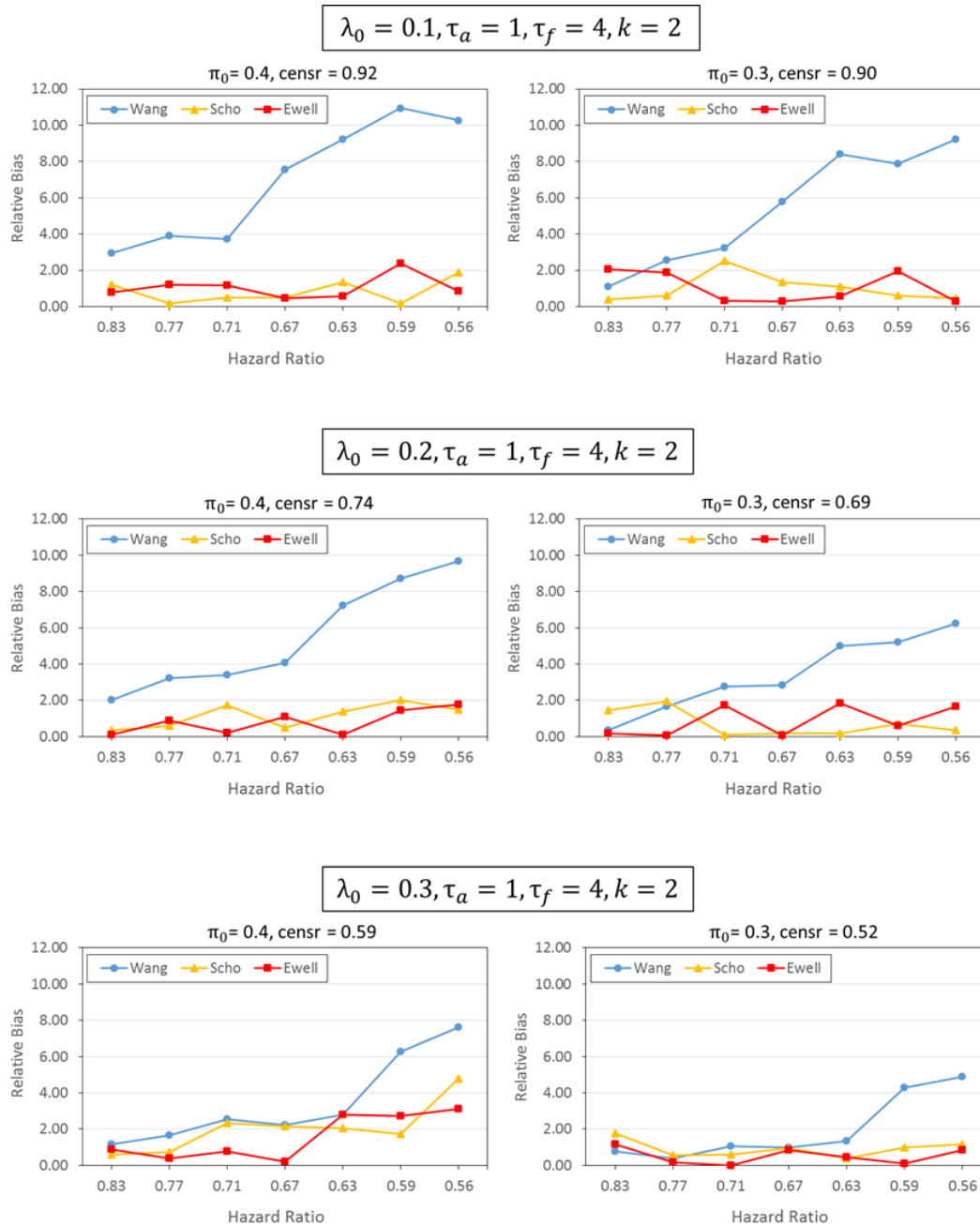
Figure 3.3 Simulation for Weibull distribution

τ_f is also changed for simulation in Weibull distribution. Similar to in the

exponential distribution, by increasing τ_f, more events could be detected, and thus lower

27

censoring rate could be achieved. By adjusting τ_f in the range of $\tau_f = 4, 10, 15, 20$, actual

mean censoring rate is achieved at $0.9 \sim 0.5$. As shown from the result in Figure 3.4,

Schoenfeld and Ewell method performs significantly better than Wang method when τ_f is

4 and 10, corresponding to the mean censoring rate 0.92 and 0.69. While when mean

censoring rate falls below 0.56, it can't be distinguished clearly among all three methods.

All of the methods are considered to perform well, as shown from that the empirical powers

are all very close to the nominal power 0.9, and the differences are all less than 4%.
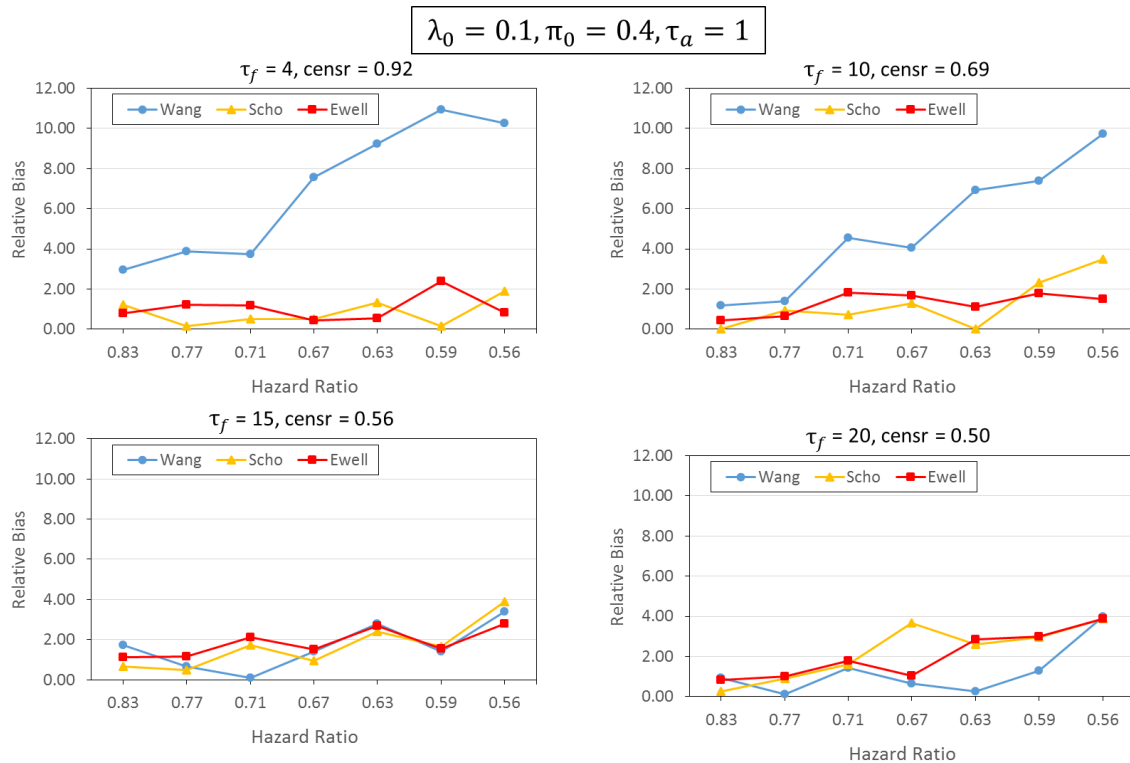
Figure 3.4 Simulation for Weibull distribution, adjusting follow up period

In lognormal distribution, mean survival time is calculated as $e^{\mu+(\sigma_0)^2/2}$. So in

contrast to the Exponential and Weibull distribution, increasing scale parameter will result

in increasing mean survival time. While keeping all other parameters constant, increasing

mean survival time will introduce higher censoring rate. The scale parameter is varied at

28
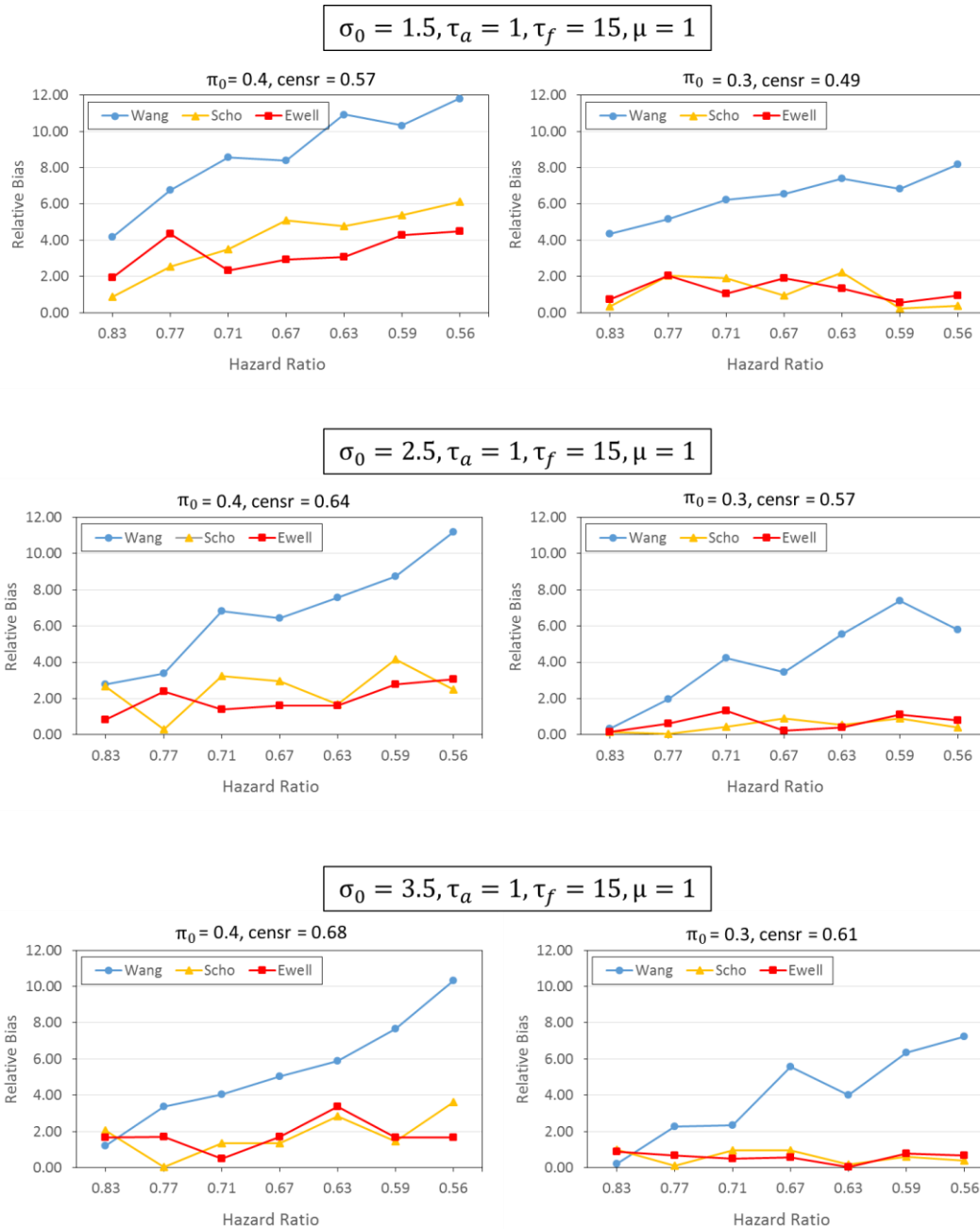
Figure 3.5 Simulation for Lognormal distribution

$scale = 1.5, 2.5, 3.5$, and the mean censoring rate is achieved at $0.48 \sim 0.67$. Comparing with the other two distributions, the results of the empirical powers for lognormal distribution is much less sensitive with varying scale parameter.

29

Figure 3.5 shows the result for lognormal distribution. For the settings used here, Schoenfeld and Ewell methods both have much better estimation capability than the Wang method in all scenarios. And we can't distinguish between the Schoenfeld and Ewell methods. Their estimations are very close.

$$\sigma_0 = 2, \pi_0 = 0.4, \tau_a = 1, \mu=1$$
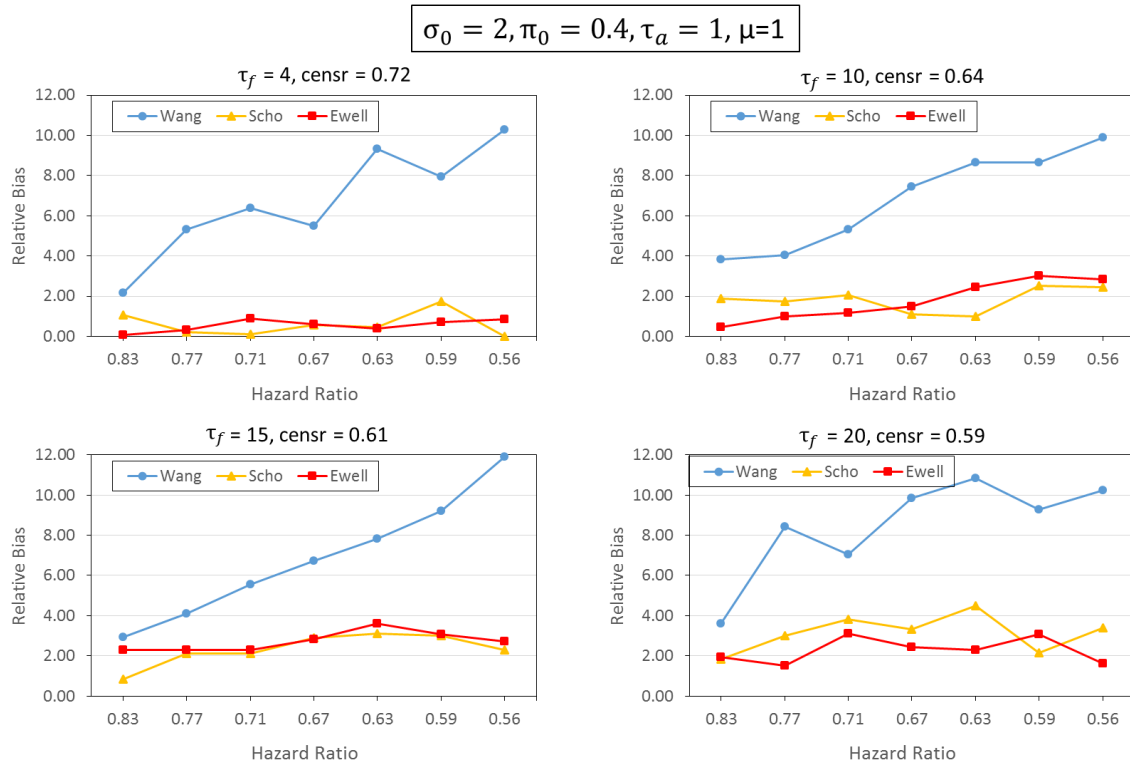
Figure 3.6 Simulation for Lognormal distribution, adjusting follow up period

When adjusting τ_f to compare the estimation performance for all three methods for lognormal distribution, similar trend is attained as Exponential and Weibull distribution. As shown in Figure 3.6, estimated empirical power from Schoenfeld and Ewell methods are much more accurate than that from Wang method, which is shown as much closer to nominal power 0.9.

3.2 COMPARISON OF PARAMETRIC SAMPLE SIZE ESTIMATION WITH NONPARAMETRIC SAMPLE SIZE ESTIMATION

3.2.1 SIMULATION SETTINGS

The model description for each setting are similar as in 3.1. The following settings are used in the simulation study for exponential and Weibull distribution:

1) scale parameters $\lambda_0 = (0.2, 0.3, 0.4, 0.5)$ and shape parameter $k = 2$ for Weibull distribution;

2) an accrual period of 3 years and a follow-up time of 4 years;

3) an equal allocation between treatment and control group, $p = 0.5$;

4) a number of 500 observations is generated in each dataset and simulation are based on 200 replication.

The settings used in the simulation study for lognormal distribution are listed as follows,

1) scale parameter $\sigma_0 = (2, 3, 4)$, shape parameter $\mu = 1$;

2) an accrual period of 1 year and a follow up period of 15 years;

3) an equal allocation between treatment and control group, $p = 0.5$;

4) a number of 500 observations is generated for each dataset and simulation results are based on 200 replications.

3.2.2 SIMULATION RESULTS

As shown in Table 3.1, for all three methods, parametric sample size, nonparametric sample size and 95% CI are reported for each set of parameters. For all the

31

three methods, the results from the nonparametric sample size estimation are very close to the results from the parametric approach, indicating that the estimation method are valid.

Similar to the exponential distribution, in Weibull and lognormal distribution, for all three methods, the results from the nonparametric sample size estimation are very close to that from the parametric approach, as shown in Table 3.2 and Table 3.3.

3.3 SIMULATION COMPARING WANG AND EWELL METHODS, FOR NONBINARY COVARIATE

3.3.1 SIMULATION SETTINGS

The following settings are used in the simulation study for Exponential distribution. Covariate data are generated from a normal distribution, with mean $\mu = 0.5$ and variance $\sigma^2 = 1$.

We specifically consider the population follows the mixture cure model with the Exponential distribution, which can be written as

$$S_0(t) = \pi_0 + (1 - \pi_0)e^{-\lambda_0 t}$$

$$\bar{S}(t) = \exp\left(-\lambda_0 t \, e^{\beta \bar{x}}\right)$$

1) an accrual period of 1 year and a follow up time of 4 years;

2) Baseline scale parameter is set at $\lambda_0 = 0.1, 0.25, 0.75.$, which corresponding to the mean survival time of baseline as 10, 4, 1.33.

3) For each λ_0, cure rate for baseline π_0 is set at 0.2 and 0.1.

4) $e^{-\beta_0}$ is at (1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8), e^{γ_0} is at (0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0).

Table 3.1 Comparison of Exponential parametric sample size estimation with nonparametric estimation

π_0	OR	λ_0	HR	Wang			Schoenfeld			Ewell		
				para	nonpara	95% CI	para	nonpara	95% CI	para	nonpara	95% CI
censoring rate: 0.66 ~ 0.75												
0.4	2.226	0.2	0.63	200	247	(83 , 581)	228	241	(84 , 554)	229	241	(84 , 553)
	2.46	0.2	0.59	158	179	(81, 408)	183	176	(79, 406)	184	176	(79, 406)
	2.718	0.2	0.56	128	142	(68, 303)	151	140	(68, 306)	152	140	(69, 306)
	2.226	0.3	0.63	171	209	(89 , 563)	185	203	(88 , 504)	185	204	(88 , 502)
	2.46	0.3	0.59	135	152	(63 , 318)	148	150	(65 , 307)	148	150	(65 , 307)
	2.718	0.3	0.56	109	128	(62 , 276)	122	127	(63 , 271)	122	127	(63 , 270)
censoring rate: 0.45 ~ 0.64												
0.4	2.226	0.4	0.63	162	190	(84 , 367)	169	187	(84 , 354)	168	187	(84 , 354)
	2.46	0.4	0.59	128	144	(76 , 250)	134	142	(76 , 249)	134	143	(76 , 249)
	2.718	0.4	0.56	104	116	(64 , 208)	110	115	(62 , 202)	110	115	(63 , 203)
0.1	2.46	0.3	0.59	164	208	(104 , 391)	186	206	(103 , 381)	186	206	(104 , 383)
	2.718	0.3	0.56	133	177	(92 , 358)	154	176	(92 , 355)	154	176	(93 , 354)
	2.718	0.2	0.56	164	209	(93 , 445)	197	206	(93 , 424)	199	206	(93 , 422)
censoring rate: 0.30 ~ 0.38												
0.1	2.226	0.4	0.63	189	255	(139 , 553)	205	252	(142 , 540)	204	252	(142 , 540)
	2.46	0.4	0.59	149	202	(100 , 379)	163	201	(102 , 377)	162	201	(101 , 379)
	2.718	0.4	0.56	121	164	(98 , 285)	134	164	(97 , 285)	134	164	(97 , 285)
	2.226	0.5	0.63	183	238	(128 , 485)	191	237	(122 , 483)	190	237	(122 , 483)
	2.46	0.5	0.59	144	194	(95 , 367)	151	192	(95 , 369)	150	192	(95 , 369)
	2.718	0.5	0.56	117	159	(89 , 283)	124	159	(90 , 287)	123	159	(90 , 287)

Table 3.2 Comparison of Weibull parametric sample size estimation with nonparametric estimation

π_0	OR	λ_0	HR	Wang para			Schoenfeld			Ewell		
				para	nonpara	95% CI	para	nonpara	95% CI	para	nonpara	95% CI
censoring rate: 0.61 ~ 0.77												
0.4	0.8	0.2	1.6	213	263	(102, 606)	243	259	(99, 593)	245	259	(99, 592)
	0.9	0.2	1.7	167	196	(79, 479)	195	190	(77, 460)	197	189	(78, 452)
	1	0.2	1.8	136	160	(72, 396)	161	155	(73, 400)	163	155	(73, 400)
	0.8	0.3	1.6	163	195	(89, 486)	169	192	(88, 465)	169	193	(89, 468)
	0.9	0.3	1.7	128	154	(75, 305)	134	152	(74, 293)	134	152	(74, 292)
	1	0.3	1.8	104	124	(62, 221)	110	122	(60, 215)	110	122	(61, 215)
censoring rate: 0.37 ~ 0.56												
0.4	0.9	0.5	1.7	148	165	(82, 322)	138	156	(78, 289)	141	159	(81, 299)
	1	0.5	1.8	120	126	(70, 226)	111	119	(66, 205)	114	121	(66, 215)
	0.9	0.4	1.7	136	150	(86, 276)	129	146	(84, 268)	130	147	(84, 272)
	1	0.4	1.8	110	120	(69, 211)	104	115	(69, 197)	105	116	(69, 199)
0.1	0.9	0.3	1.7	149	205	(109, 394)	163	206	(110, 396)	162	206	(110, 396)
	1	0.3	1.8	121	179	(92, 400)	134	181	(94, 397)	133	181	(93, 397)
censoring rate: 0.17 ~ 0.25												
0.1	0.8	0.4	1.6	187	231	(117, 463)	183	228	(116, 465)	183	228	(116, 466)
	0.9	0.4	1.7	147	187	(103, 364)	144	184	(99, 365)	144	184	(99, 367)
	1	0.4	1.8	120	149	(82, 259)	117	147	(79, 248)	117	148	(80, 248)
	0.8	0.5	1.6	203	237	(116, 497)	191	226	(109, 480)	194	229	(111, 484)
	0.9	0.5	1.7	160	184	(102, 385)	149	176	(95, 369)	152	179	(97, 373)
	1	0.5	1.8	130	151	(85, 287)	121	144	(78, 276)	123	146	(81, 279)

Table 3.3 Comparison of Lognormal parametric sample size estimation with nonparametric sample size estimation

π_0	OR	σ_0	HR	Wang			Schoenfeld			Ewell		
				para	nonpara	95% CI	para	nonpara	95% CI	para	nonpara	95% CI
censoring rate: 0.49 ~ 0.59												
0.3	2.014	2	0.67	182	251	(126 , 641)	208	248	(125 , 640)	210	249	(126 , 640)
	2.226	2	0.63	137	196	(91 , 435)	160	194	(89 , 435)	162	194	(89 , 435)
	2.46	2	0.59	108	144	(76 , 263)	129	143	(76 , 263)	130	143	(76 , 263)
	2.226	3	0.63	159	196	(94 , 460)	184	195	(93 , 460)	185	195	(93 , 460)
	2.46	3	0.59	125	158	(82 , 318)	147	157	(82 , 317)	149	157	(82 , 317)
	2.226	2	0.56	88	122	(71 , 211)	107	121	(69 , 210)	108	121	(70 , 210)
censoring rate: 0.61 ~ 0.65												
0.3	2.718	3	0.56	102	129	(65 , 277)	122	128	(65 , 277)	123	128	(65 , 277)
	2.226	4	0.63	174	194	(91 , 416)	200	193	(91 , 415)	200	193	(91 , 415)
	2.46	4	0.59	137	155	(76 , 310)	160	154	(76 , 308)	161	154	(76 , 309)
	2.718	4	0.56	112	126	(68 , 243)	133	125	(68 , 241)	133	125	(68 , 242)
0.4	2.46	2	0.59	104	142	(77 , 277)	124	141	(76 , 276)	125	141	(77 , 276)
	2.718	2	0.56	85	115	(66 , 218)	102	114	(65 , 214)	104	114	(65 , 215)
censoring rate: 0.66 ~ 0.71												
0.4	2.226	3	0.63	154	188	(84 , 345)	177	187	(84 , 345)	178	187	(84 , 345)
	2.46	3	0.59	121	145	(72 , 288)	142	144	(72 , 287)	143	144	(72 , 287)
	2.718	3	0.56	99	114	(61 , 205)	118	113	(61 , 205)	119	113	(61 , 205)
	2.226	4	0.63	170	202	(89 , 422)	194	201	(88 , 418)	194	201	(88 , 418)
	2.46	4	0.59	134	167	(77 , 370)	155	165	(76 , 368)	156	165	(76 , 368)
	2.718	4	0.56	109	114	(62 , 217)	128	114	(62 , 217)	129	114	(62 , 217)

Settings used in the simulation study for Weibull distribution are listed as follows

The population follows the mixture cure model with the Weibull distribution, which can be written as

$$S_0(t) = \pi_0 + (1 - \pi_0)e^{-(\lambda_0 t)^k}$$

$$\bar{S}(t) = \exp\left(-(\lambda_0 t)^k e^{\beta \bar{x}}\right)$$

1) an accrual period of 1 year and a follow up time of 4 years;

2) Baseline scale parameter is set at $\lambda_0 = 0.1, 0.2, 0.5.$, which corresponding to the mean survival time of baseline as 10, 5, 2. Shape parameter is 2.

3) For each λ_0, cure rate for control arm π_0 is set at 0.2 and 0.1.

4) $e^{-\beta_0}$ is at (1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8), e^{γ_0} is at (0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0).

Settings used in the simulation study for lognormal distribution are listed as follows:

The population follows the mixture cure model with the lognormal distribution, which can be written as

$$S_0(t) = \pi_0 + (1 - \pi_0)(1 - pnorm\left(\frac{lnt - \mu_0}{\sigma_0}\right))$$

$$\bar{S}(t) = \exp\left(-(\lambda_0 t)^k e^{\beta \bar{x}}\right)$$

1) an accrual period of 1 year and a follow up time of 20 years;

2) Baseline scale parameter is set at $\sigma_0 = 2, 2.5, 3.5.$, which corresponding to the mean survival time of baseline as 5, 4, 0.286. Shape is 0.1.

3) For each λ_0, cure rate for baseline π_0 is set at 0.4 and 0.5.

4) $e^{-\beta_0}$ is at (1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8), e^{γ_0} is at (0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0).

Effects of follow up period are investigated on the Weibull distribution as an example, the settings are listed as follows:

36

1) an accrual period of 1 year and baseline scale parameter is set at $\lambda_0 = 2$,

2) follow up period is set at $\tau_f = 5, 10, 15, 25$.

3) π_0 is set at 0.4, and shape parameter $k = 0.1$.

4) $e^{-\beta_0}$ is at $(1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8)$, e^{γ_0} is at $(0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0)$.

3.3.2 SIMULATION RESULTS

For the exponential distribution, the mean survival time is calculated as $^1/_{\lambda_0}$. Corresponding to $\lambda_0 = (0.1, 0.25, 0.75)$, mean censoring rate is attained at $(0.82, 0.64, 0.41)$ for $\pi_0 = 0.2$, and at $(0.78, 0.56, 0.28)$ for $\pi_0 = 0.1$. As shown in Figure 3.7, if looking at the left upper panel for case of relatively high censoring rate (censr = 0.82), the empirical power calculated by Wang's method significantly different from the nominal power. The Ewell method worked very well, in which the largest deviation is only 4% from the nominal power. With the censoring rate decreasing to 0.64 and 0.41, Wang method performs better, while the Ewell maintains good performance in all scenarios. In censoring rate of 0.41, all the cases got good estimation, with nearly all deviation of empirical power from nominal power smaller than 3%. Similar trend has been observed for the three graphs in right panel, where $\pi_0 = 0.1$.

For the Weibull distribution, mean survival time is calculated as $\frac{\Gamma(1+1/k)}{\lambda_0}$. As the result, larger λ_0 will lead to shorter mean survival time, and thus higher censoring rate. For $\pi_0 = 0.2$, corresponding to $\lambda_0 = (0.1, 0.25, 0.75)$, mean censoring rate is achieved at $(0.91, 0.71, 0.37)$. Similar trend is observed for Weibull distribution, as for exponential distribution in Figure 3.8. The Ewell method shows superior performance over the Wang

37

method in relative high and medium censoring rate (censr = 0.91 and 0.71). While for low censoring rate (censr = 0.37), both of the methods perform well, nearly all the deviation of empirical power from the nominal power is less than 2%. Similar trend is observed for $\pi_0 = 0.1$. In which, the mean censoring rate is (0.89, 0.65, 0.23).
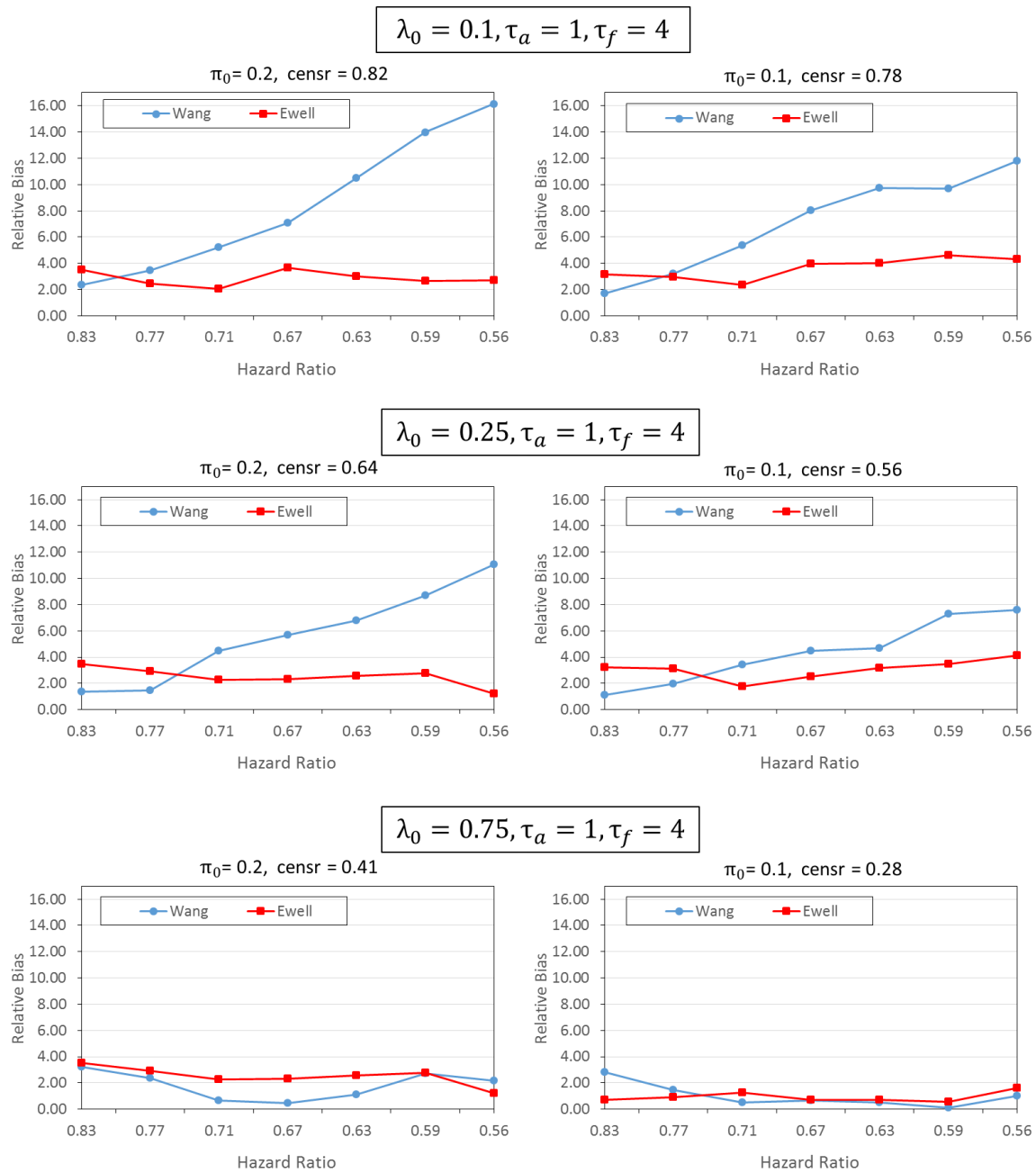

Figure 3.7 Comparing two methods for exponential distribution, for nonbinary covariate

Effects of follow up period τ_f is studied for Weibull distribution. By adjusting τ_f in the range of $(4, 10, 15, 25)$, actual mean censoring rate is achieved at $(0.91, 0.64, 0.47, 0.36)$. Similar as the case for Weibull distribution in binary case, by increasing τ_f, more events could be detected, and thus lower censoring rate will be achieved.
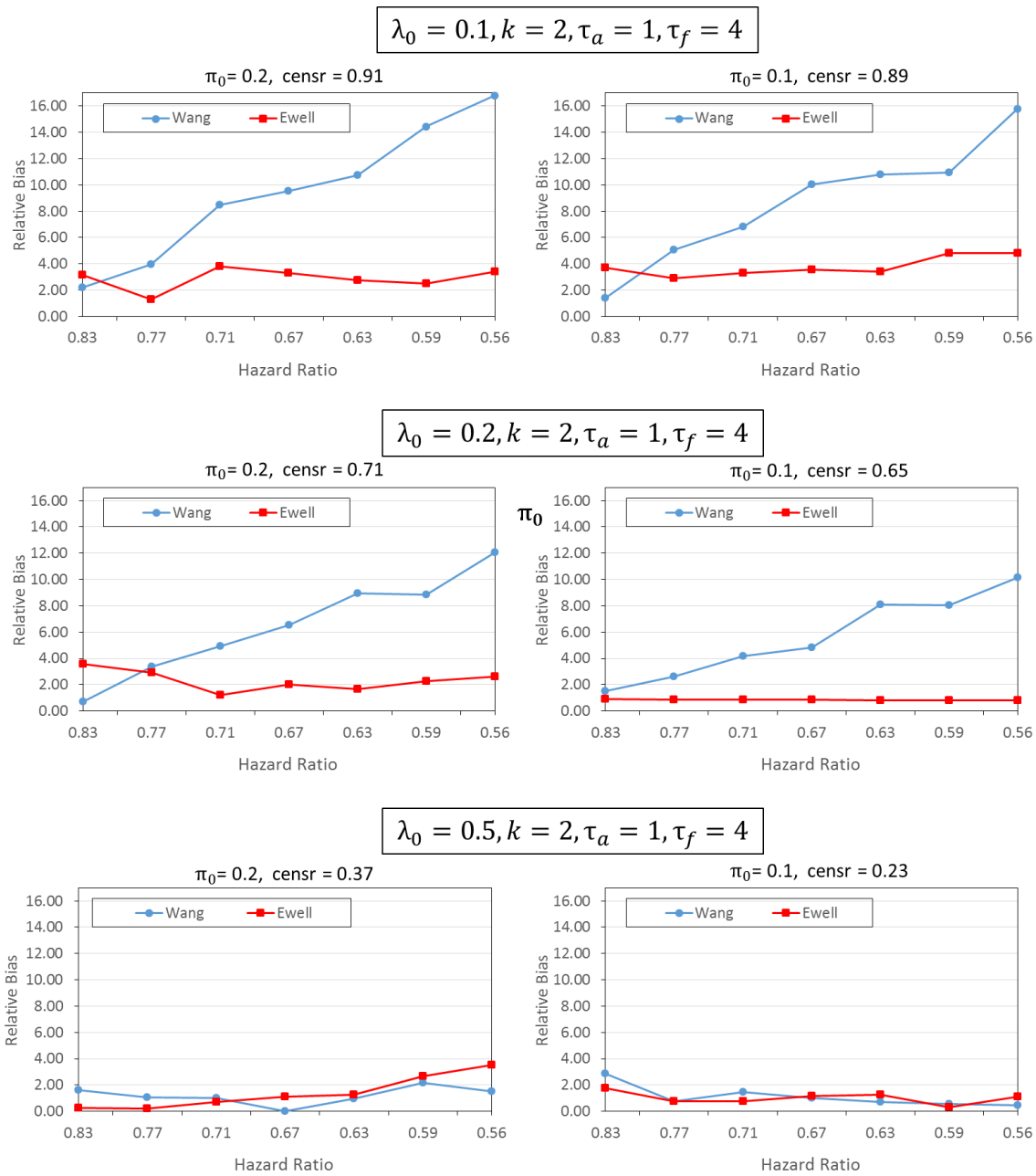

Figure 3.8 Comparing two methods for Weibull distribution, for nonbinary covariate

As shown in Figure 3.9, the Ewell method performs significantly better than the Wang method when τ_f is 4 and 10. When extending τ_f to 15, Ewell method becomes only slightly better than Wang method. When τ_f reaches 25, corresponding to a censoring rate of 0.36, we couldn't distinguish between the two methods, with all the deviation of empirical power from the nominal power less than 4%.

Different from the exponential and Weibull distribution, increasing λ_0 results in increasing censoring rate in lognormal distribution, as shown in Figure 3.10. For $\pi_0 = 0.4$, by varying σ_0 at (2, 2.5, 3.5), mean censoring rate is acquired at (0.63, 0.66, 0.70). As seen from here, censoring rate is relatively stable while changing σ_0. With all the settings here, Ewell method performs significantly better than the Wang method. Similar trend is observed for $\pi_0 = 0.5$ as for $\pi_0 = 0.4$.

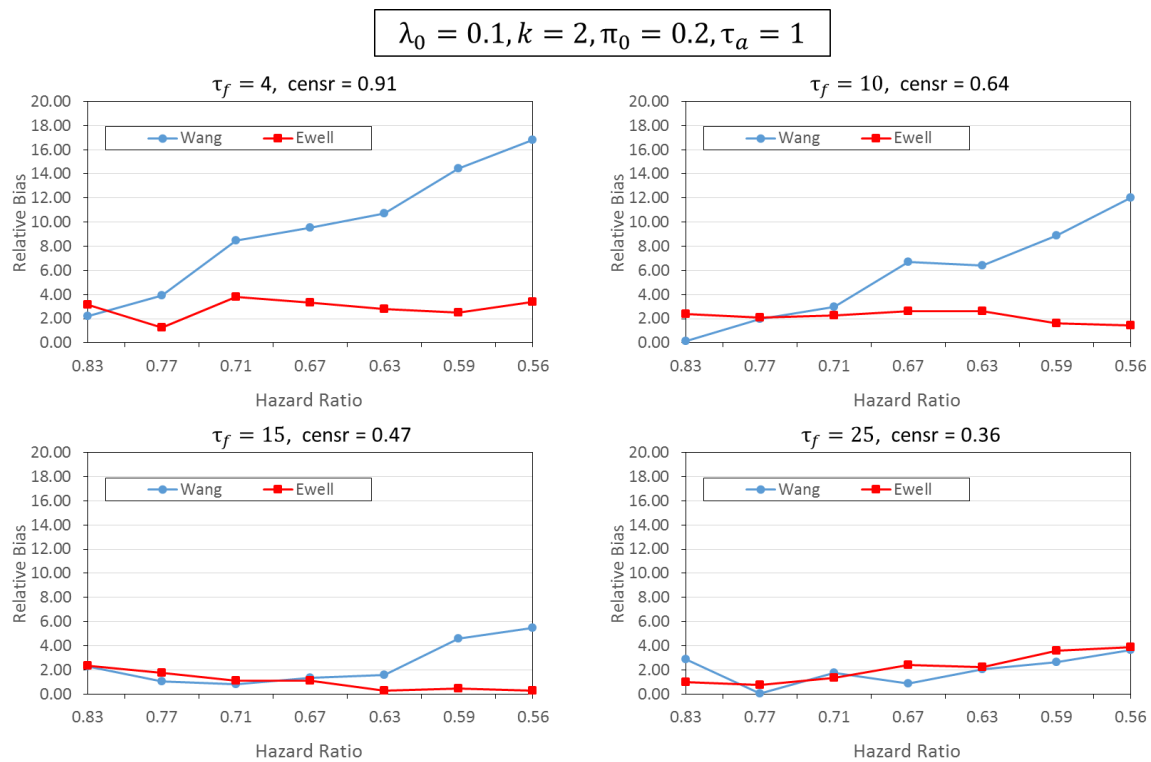$$\lambda_0 = 0.1, k = 2, \pi_0 = 0.2, \tau_a = 1$$

Figure 3.9 Comparing the Wang and Ewell methods for Weibull distribution, for nonbinary covariate, adjusting follow up period
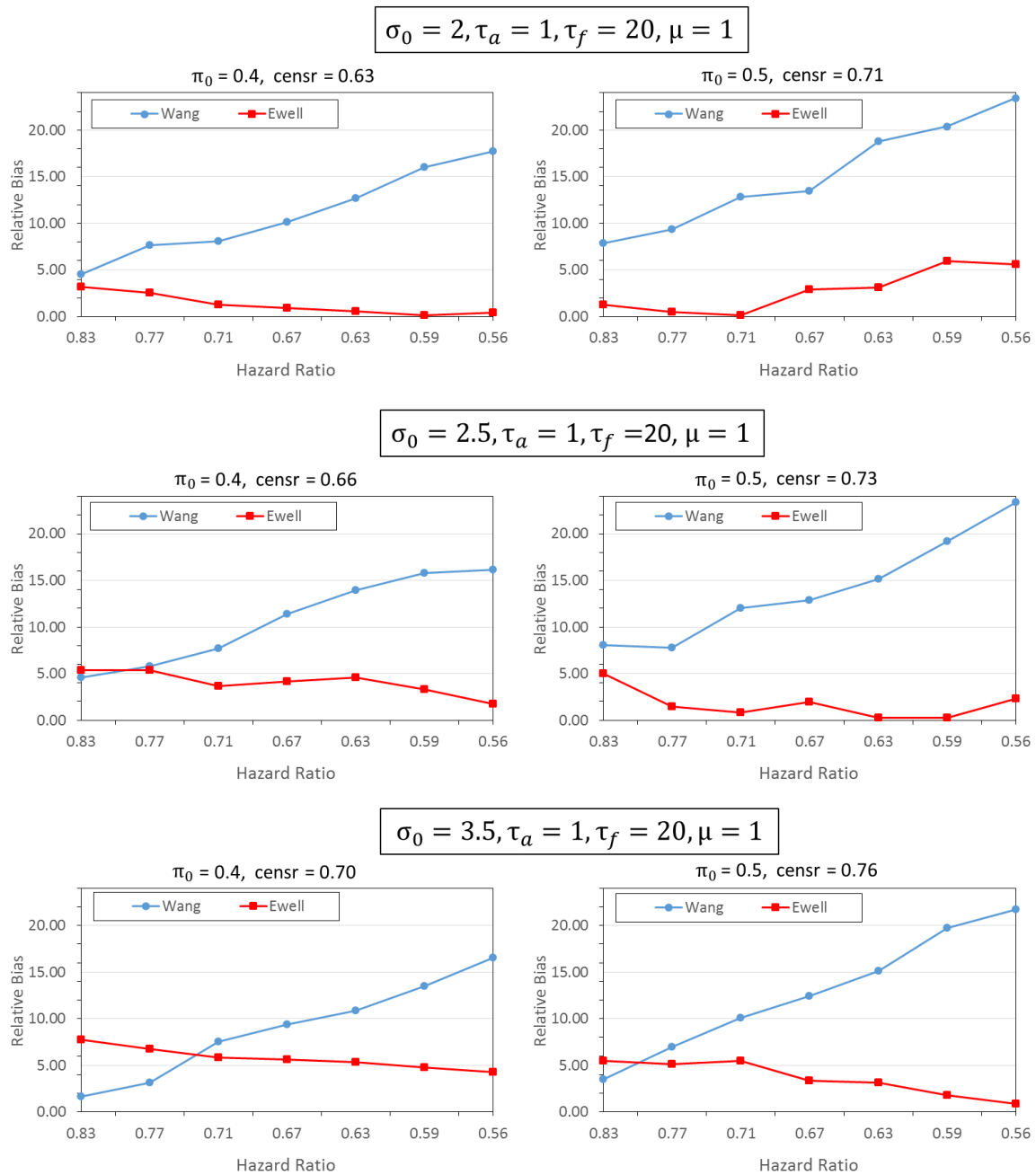
Figure 3.10 Comparing the Wang and Ewell methods for Lognormal distribution, for nonbinary covariate

CHAPTER 4 REAL DATA APPLICATION AND POWER ANALYSIS

The components needed for the sample size calculation package include: sample size (power), censoring distribution (accrual period, follow up time, accrual distribution), and components specific for the mixture cure model (hazard ratio e^{β_0}, odds ratio (e^{γ_0}, π_1, π_0), survival distribution). In our package, there are two syntaxes for specifying the survival distribution: one is parametric syntax assuming the Exponential (Weibull, or Lognormal) survival distribution with all the parameters defined; the other is nonparametric syntax using the known dataset to estimate all the necessary parameters through *smcure* package. In this chapter, we will illustrate specifically how to apply the package in both parametric and nonparametric approaches.

In section 4.1, we introduce the parametric sample size estimation by using a specific example under Weibull distribution. Both the syntax and output interface is demonstrated. Especially for output, it's shown that the function is flexible with the option "method" that can either output the result from Schoenfeld method only, or can output the results from all three methods at the same time. Section 4.2 illustrates the application of the package to a well-known data set Eastern Cooperative Oncology Group trial e1684, as well as the calculation for cure rates. Section 4.3 studies the association between the power and the sample size by power curve analysis. Section 4.4 conducts the power curve analysis under various hazards ratios, as well as the power curve analysis under various odds ratios.

4.1 PARAMETRIC SAMPLE SIZE ESTIMATION

The package assumed that both arms follow the same distribution, such as exponential, Weibull or lognormal. For example, if a survival trial follows a uniform accrual with an accrual period of 2 years and a follow-up period of 5 years, with equal allocation of patients in each arm ($p = 0.5$). The mean life of uncured patients is 3 years for control arm and 4 years for treatment arm. It's assumed that both arms follow the Weibull distribution with $k = 2$. Cure rates are $\pi_0 = 0.2$ and $\pi_1 = 0.3$ for the control and treatment arm, respectively. In order to detect a 33% improvement in the mean survival time from 3 to 4 years, at 95% significance level and 90% power, the estimated sample size can be acquired by the following code:

```
> allNPHMC(power=0.9, alpha=0.05, accrualtime=2, followuptime=5,
+ p=0.5,accrualdist="uniform",hazardratio=3/4,oddsratio=1.714,
+ pi0=0.2,data=NULL,survdist="weib", shape=2, scale=0.295)
```

The output is as follows:

```
==========================================================================
SAMPLE SIZE CALCULATION FOR PH MIXTURE CURE MODEL AND STANDARD PH MODEL
==========================================================================
At alpha= 0.05 and power= 0.9 :
            PH Mixture Cure Model:n = 411

            Standard PH Model:n = 551
```

As shown from the above result, a sample size of 411 patients will be needed to achieve a power of 90% based on PH mixture cure model. While the sample size estimated from the standard PH model is 551, which apparently underestimates the power.

If the "method" option is set as "all", results from all the three methods will be displayed together as follows,

```
> allNPHMC(power=0.9, alpha=0.05, accrualtime=2, followuptime=5,
+ p=0.5,accrualdist="uniform",hazardratio=3/4,oddsratio=1.714,
+ pi0=0.2,data=NULL,survdist="weib", shape=2, scale=0.295,method="all")

========================================================================
SAMPLE SIZE CALCULATION FOR PH MIXTURE CURE MODEL AND STANDARD PH MODEL
========================================================================
At alpha= 0.05 and power= 0.9 :
            PH Mixture Cure Model:
 Wang = 418     Schoenfeld = 411     Ewell = 412

            Standard PH Model:
 Wang = 535     Schoenfeld = 551     Ewell = 549
```

4.2 REAL DATA ANALYSIS

The application of our sample size calculation package to the real data is illustrated with the dataset from Eastern Cooperative Oncology Group trial e1684 [18]. Note that our intension here is not to re-design the trial but to show the sample size requirement under the mixture cure model when the hazard ratio and cure rates are estimated from the dataset itself. The ECOG e1684 is a two-arm phase III clinical trial comparing high-dose interferon alpha-2b with an observation arm. The primary endpoint was relapse-free survival (RFS), with RFS defined as the time from randomization until progression of the tumor or death. There are a total of 286 patients enrolled and the total study time is 7 years, with an accrual period of 4 years and a follow-up period of 3 years. As seen from the Kaplan Meier survival curve, even in the end of the study (not shown here), the survival probability can still be as high as about 0.2. So in order to account for this curable part, mixture cure model should be applied here.

To do nonparametric estimation, only *power, alpha, accrualtime, followuptime, p, accrualdist* and *data* need to be specified, with the hazard ratio and cure rates being directly estimated from the dataset by *smcure* package. The following code is used to obtain the sample size:

```
> allNPHMC(power=0.8,alpha=0.05,accrualtime=4,followuptime=3,p=0.5,
+          accrualdist="uniform",data=e1684szdata)
```

The output is as follows:

```
Program is running..be patient... done.
Call:
smcure(formula = Surv(Time, Status) ~ X, cureform = ~X, data = data,
    model = "ph", Var = FALSE)

Cure probability model:
              Estimate
(Intercept)  1.2850677
X           -0.5455204


Failure time distribution model:
    Estimate
X -0.1643542


=====================================================================
SAMPLE SIZE CALCULATION FOR PH MIXTURE CURE MODEL AND STANDARD PH MODEL
=====================================================================
At alpha = 0.05 and power = 0.8 :
              PH Mixture Cure Model:n = 459

  Standard PH Model with KM estimators:n = 255
```

The data are first fitted by recalling the *smcure* R package with the treatment as a covariate. The $\log(hazard\ ratio)$ β_0 is estimated as -0.164 from the *smcure* package. The coefficient from logistic regression model for modeling the cured part is 1.285 and -0.546, which results in the cure rates for the observation arm and treatment arm calculated as follows:

$$\hat{\pi}_0 = 1 - e^{1.285}/(1 + e^{1.285}) = 0.217$$

$$\hat{\pi}_1 = 1 - (e^{1.285-0.546})/(1 + e^{1.285-0.546}) = 0.323$$

And thus to achieve 80% power, a sample size of 459 is required based on the estimates from the PH mixture cure model, while only 255 is required if ignoring the cured part. As is clearly seen, under this set of parameters, the cure fraction will lead to an underpowered trial.

4.3 POWER ANALYSIS

Besides the sample size calculation, the package can provide power analysis if the sample sizes are given. Using the same setting as in 4.1, the power of the sample size of 100, 150, 200, 250, 300, 350, 400, 450 and 500, can be obtained by using the following code:

```
> n=seq(100,500,by=50)
> allNPHMC(n=n, alpha=0.05, accrualtime=2, followuptime=5, p = 0.5,
+ accrualdist ="uniform", hazardratio = 3/4, oddsratio = 1.714, pi0 = 0.2,
+ data = NULL, survdist = "weib", shape = 2, scale = 0.295)
```

```
=====================================================================
POWER CALCULATION FOR PH MIXTURE CURE MODEL AND STANDARD PH MODEL
=====================================================================
         PH Mixture Cure Model:power = 0.36 0.5 0.62 0.72 0.79 0.85 0.89 0.92 0.95

         Standard PH Model:power = 0.28 0.39 0.5 0.59 0.67 0.73 0.79 0.83 0.87
```

As seen from the Figure 4.1, with the increase of sample size from 100 to 500, the power will increase from 0.36 to 0.95 for PH mixture cure model, and from 0.28 to 0.87 for standard PH model. Comparing the results between PH mixture cure model and standard PH model, the power is underestimated with ignoring the cure fraction.

Figure 4.2. shows how to calculate the power using the e1684 data. For both of the models, the calculated power increase with increasing sample size. However, ignoring the cure fraction leads to overestimating the power.

4.4 POWER ANALYSIS FOR STUDYING EFFECTS OF HAZARDRATIO AND ODDSRATIO

Power analysis is conducted to study the effects of the hazard ratio and odds ratio. For example, a survival trial will follow a uniform accrual with an accrual period of 3 years and a follow-up period of 4 years with equal allocation of patients in each arm ($p = 0.5$).
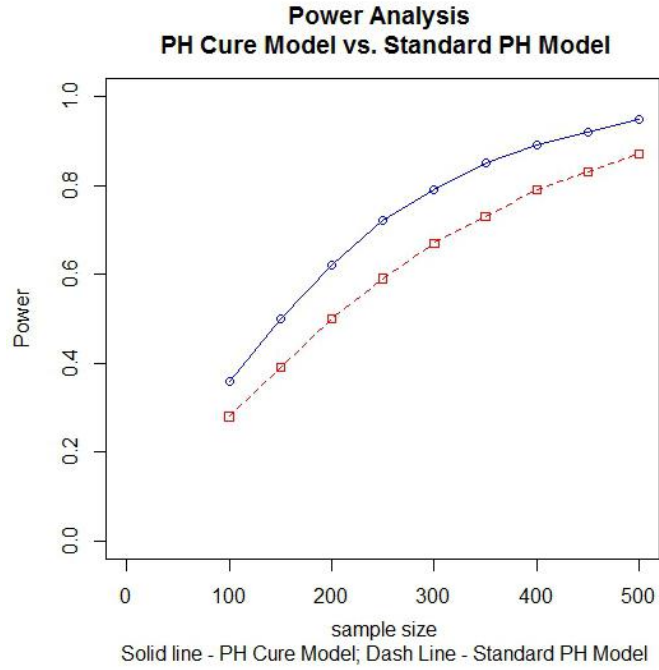
46

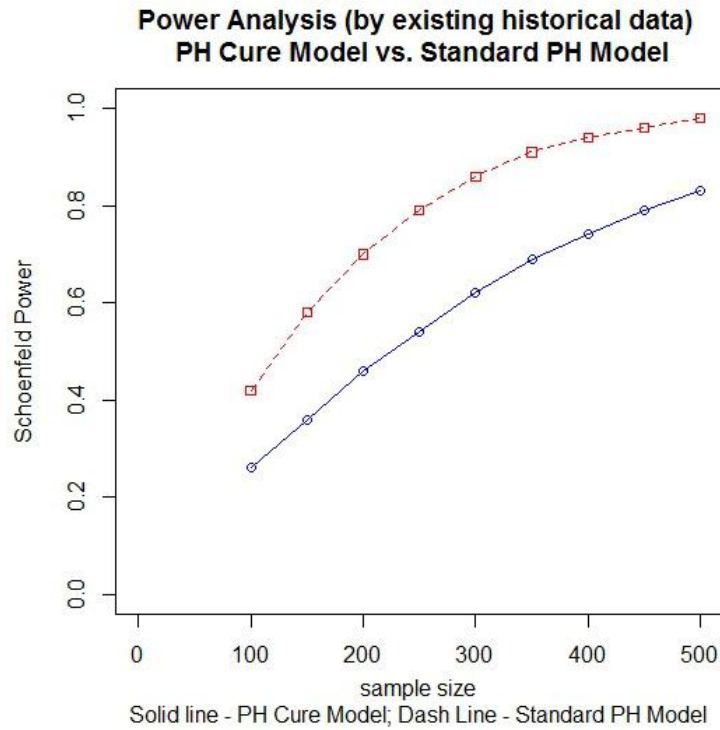Figure 4.1 Power analysis for parametric estimation

Figure 4.2 Power analysis for nonparametric estimation

47

The mean life of uncured patients in the control arm will be 2 years. Both arms follow the Exponential distribution and cure rate for control arm is 0.1. Oddsratio is set at 1.5. Hazard ratio is varied at (0.56, 0.59, 0.63, 0.67, 0.71, 0.77, 0.83). Powers under different sample sizes (100, 200, 300, 400, 500) are plotted in Figure 4.3. Generally, power decreases with increasing hazard ratio. Higher hazard ratio indicates smaller difference between two arms' uncured parts. And thus it will result in the lower power in detecting the difference between the two arms. With the same hazard ratio, power will increase with the increase of the sample size.
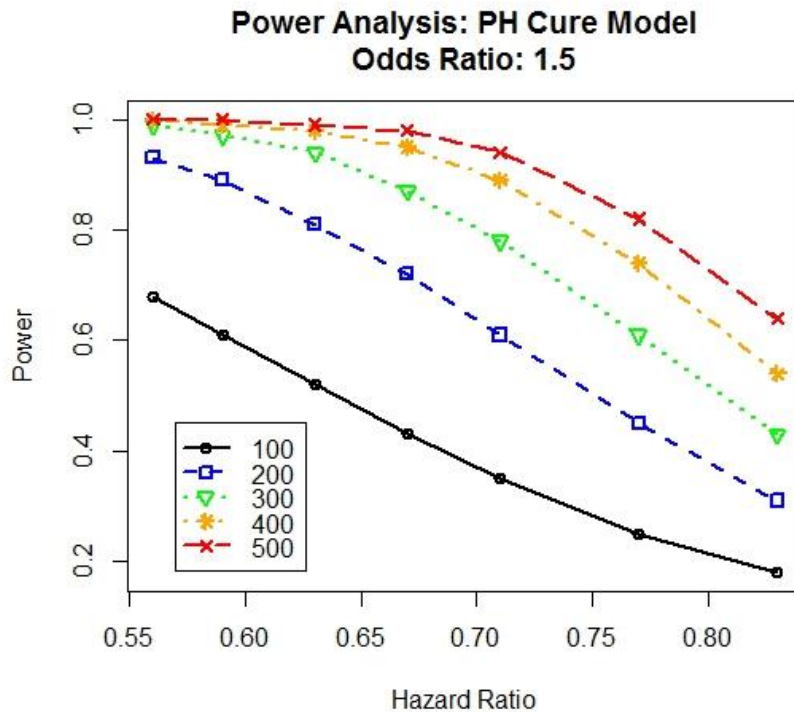

Figure 4.3 Effects of hazardratio on power under different sample sizes

The same setting is used to study the effects of oddsratio on power. The hazard ratio is set at 0.83, while the oddsratio is varied at (2.718, 2.460, 2.226, 2.014, 1.822, 1.649, 1.492). Similarly, powers under different sample sizes (100, 200, 300, 400, 500) are plotted

48

in Figure 4.4. Power increases with increasing odds ratio. Larger odds ratio means more significant differences between the cure rates of two arms, and thus easier to be differentiated. With the same odds ratio, increasing sample size will increase the power.
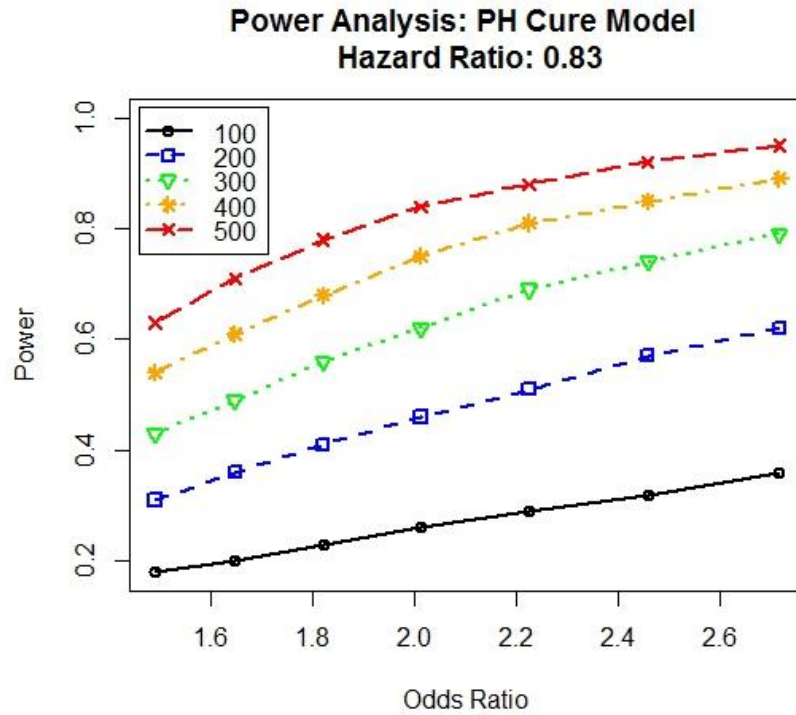

Figure 4.4 Effects of odds ratio on power under different sample sizes

CHAPTER 5 SUMMARY AND CONCLUSIONS

The development of modern medical technology has substantially improved cure rates for many fatal and chronic diseases. The mixture cure model is more appropriate to be applied to this type of study in order to account for the cured fractions among patients. The two components of the mixture cure model can model the cured and uncured proportions separately and simultaneously.

Sample size calculation is crucial for designing randomized controlled clinical trials in survival analysis. The current sample size calculation method for PH mixture cure model naturally extends the sample size formula for standard PH model by allowing a cure fraction. However, this formula has limitations in calculating the probability of death, which is currently contributed only from the control arm, and this may underestimate or overestimate the required sample size. In the thesis, we have investigated the two modified versions of the sample size calculations for the PH mixture cure model. First we propose two methods for correcting the calculation of probability of death, by acquiring the numeric mean and geometric mean of the probability of death based on both control arm and treatment arm. Extensive simulation study has shown that the two proposed methods performs significantly better than the WANG's method under various scenarios, when the study has the relatively high and medium censoring rate. There're no significant differences between the two proposed methods.

Secondly, the method proposed by Wang has the parametric assumption on survival function under the exponential and Weibull distribution. We extend all three methods to the cases with the survival function under lognormal distribution. A simulation study has illustrated that the two proposed methods performs much better than the Wang method in the scenarios of relative high and medium censoring rate under lognormal distribution, similar as exponential and Weibull distribution. When the censoring rate is low, all three methods has good estimation for the required sample size. All the above proposed methods and extensions have been implemented in the R package. The method has been further applied to the melanoma dataset from the ECOG phase III clinical trial e1684. The power curve analysis has been conducted to demonstrate the application of the proposed methods.

Thirdly, the sample size calculation for PH mixture cure model has been extended to the nonbinary (continuous) covariate cases under Wang method and Ewell method. The performances based on two approaches are evaluated by the comprehensive simulations and we also observe the preferred performance under the Ewell method. The function has been developed under this case and more verification needed before releasing this function as a R package.

In the future, there are many potential extensions and research directions for the sample size calculation under the mixture cure models. For example, it is common to see more than one treatment option in clinical trial. The sample size formula has been extended to the multi-arm cases [19] and it is possible to adapt similar approach to the PH mixture cure model for multi-arm covariates. Other challenging and interesting topics include developing the sample size calculation under the time varying coefficient structure. There are some other common scenarios seen in clinical trial, such as loss to follow-up and

51

noncompliance such as cross over effects [20, 21], which can also be considered applied to the sample size calculation in PH mixture cure model to improve the accuracy of sample size and power calculation.

BIBLIOGRAPHY

1. Breslow, N.E., *Analysis of Survival Data under the Proportional Hazards Model.* International Statistical Review / Revue Internationale de Statistique, 1975. **43**(1): p. 45-57.

2. Cox, D.R., *Regression Models and Life-Tables.* Journal of the Royal Statistical Society. Series B (Methodological), 1972. **34**(2): p. 187-220.

3. Andersen, P.K. and R.D. Gill, *Cox's Regression Model for Counting Processes: A Large Sample Study.* The Annals of Statistics, 1982. **10**(4): p. 1100-1120.

4. Schoenfeld, D., *The Asymptotic Properties Of Nonparametric-Tests For Comparing Survival Distributions.* Biometrika, 1981. **68**(1): p. 316-319.

5. Schoenfeld, D.A., *Sample-Size Formula For The Proportional-Hazards Regression-Model.* Biometrics, 1983. **39**(2): p. 499-503.

6. Institute, S., *SAS/STAT User's guide: Version 8, volume 1.* 1999.

7. Weiliang Qiu, J.C., Ross Lazarus, Bernard Rosner, and Jing Ma., *Powersurvepi: Power and Sample Size Calculation For Survival Analysis Of Epidemiological Studies.* R package version 0.0.5, 2009.

8. Kuk, A.Y.C. and C.H. Chen, *A Mixture Model Combining Logistic-Regression With Proportional Hazards Regression.* Biometrika, 1992. **79**(3): p. 531-541.

9. Peng, Y.W. and K.B.G. Dear, *A Nonparametric Mixture Model For Cure Rate Estimation.* Biometrics, 2000. **56**(1): p. 237-243.

10. Sy, J.P. and J.M.G. Taylor, *Estimation In a Cox Proportional Hazards Cure Model.* Biometrics, 2000. **56**(1): p. 227-236.

11. Boag, J.W., *Maximum Likelihood Estimates Of The Proportion Of Patients Cured By Cancer Therapy.* Journal of the Royal Statistical Society Series B-Statistical Methodology, 1949. **11**(1): p. 15-53.

12. Cai, C., et al., *smcure: An R-package for Estimating Semiparametric Mixture Cure Models.* Computer methods and programs in biomedicine, 2012. **108**(3): p. 1255-1260.

13. Corbiere, F. and P. Joly, *A SAS Macro For Parametric And Semiparametric Mixture Cure Models.* Computer Methods and Programs in Biomedicine, 2007. **85**(2): p. 173-180.

14. Wang, S.F., J.J. Zhang, and W.B. Lu, *Sample Size Calculation For The Proportional Hazards Cure Model.* Statistics in Medicine, 2012. **31**(29): p. 3959-3971.

15. Cai, C., et al., *NPHMC: An R-Package for Estimating Sample Size Of Proportional Hazards Mixture Cure Model.* Computer Methods and Programs in Biomedicine, 2014. **113**(1): p. 290-300.

16. Ewell, M. and J.G. Ibrahim, *The Large Sample Distribution of The Weighted Log Rank Statistic Under General Local Alternatives.* Lifetime data analysis, 1997. **3**(1): p. 5-12.

17. Hsieh, F.Y. and P.W. Lavori, *Sample-Size Calculations for The Cox Proportional Hazards Regression Model With Nonbinary Covariates.* Controlled Clinical Trials, 2000. **21**(6): p. 552-560.

18. Kirkwood, J.M., et al., *Interferon alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: The Eastern Cooperative Oncology Group trial EST 1684.* Journal of Clinical Oncology, 1996. **14**(1): p. 7-17.

19. Jung, S.-H., C. Kim, and S.-C. Chow, *Sample Size Calculation for The Log-Rank Tests for Multi-Arm Trials with a Control.* Journal of the Korean Statistical Society, 2008. **37**(1): p. 11-22.

20. Barthel, F.M.S., et al., *Evaluation of Sample Size and Power for Multi-Arm Survival Trials Allowing for Non-Uniform Accrual, Non-Proportional Hazards, Loss To Follow-Up and Cross-Over.* Statistics in Medicine, 2006. **25**(15): p. 2521-2542.

21. Lu, Q.S., et al., *Analysis of Time-To-Event Data With Nonuniform Patient Entry And Loss To Follow-Up under A Two-Stage Seamless Adaptive Design With Weibull Distribution.* Journal of Biopharmaceutical Statistics, 2012. **22**(4): p. 773-784.