

Analytical Strategies for Failure Time Data with a Cured Fraction

Sheng-li An (✉ asl0418@126.com)

Southern Medical University <https://orcid.org/0000-0002-2811-7759>

Fuqiang Huang

Southern Medical University

Pei Kang

Southern Medical University

Yingxin Liu

Southern Medical University

Fu-qiang Huang

sourtren medical university

Ying-xin Liu

sourthren medical university

Research article

Keywords: Cox's PH model, PHMC model, AFT model, AFTMC model, cure model

Posted Date: January 6th, 2020

DOI: <https://doi.org/10.21203/rs.2.18138/v2>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Analytical Strategies for Failure Time Data with a Cured Fraction

Sheng-li An^{1#}, Fu-qiang Huang^{1#}, Pei Kang¹, Ying-xin Liu¹

Authors' Affiliations: ¹State Key Laboratory of Organ Failure Research, National Clinical Research Center for Kidney Disease, Guangzhou, China; Department of Biostatistics, School of Public Health, Southern Medical University, Guangzhou, China

#Sheng-li An and Fu-qiang Huang contributed equally to this work.

Corresponding Authors: Sheng-li An, State Key Laboratory of Organ Failure Research, National Clinical Research Center for Kidney Disease, Department of Biostatistics, School of Public Health, Southern Medical University, 1838 North Guangzhou Avenue, Guangzhou 510515, China. Phone: 020-61649465; Email: asl0418@126.com

Abstract

Background: Some failure time data comes from a population that consists of some subjects that are susceptible to and others that are non-susceptible to the event of interest. The data typically have heavy censoring at the end of the follow-up period, and a traditional survival analysis would not always be appropriate. Yet it is commonly seen in literatures.

Methods: We carry out simulation studies to compare the performances of Cox's PH model with proportional hazards mixture cure (PHMC) model and accelerated failure model (AFT model) with AFT mixture cure (AFTMC) model respectively. Then we apply the models to the datasets of Lung Cancer and Eastern Cooperative Oncology Group (ECOG) phase III clinical trial E1684.

Results: When the cured rate is 0, the estimated bias, confidence interval capture rate, and K index of PHMC and AFTMC model are close to Cox's PH and AFT model respectively. The MSE of PHMC model is slightly larger than Cox's PH model and of AFTMC model are similar to AFT model. When survival data has a substantial proportion of subjects being cured, the absolute value of Bias and MSE in PHMC and AFTMC model are always smaller than Cox's PH and AFT model respectively. The confidence interval capture rate of PHMC and AFTMC model are always closer to the acceptable range than Cox's PH and AFT model. The K index of PHMC and AFTMC model are always greater than Cox's PH and AFT model.

Conclusions: The PHMC and AFTMC model do not have obvious advantages for time-to-event

data without a cured fraction. In this case, it is recommended to utilize Cox's PH or AFT model for analysis. If some subjects are non-susceptible to the event of interest in the data, it is recommended to utilize PHMC or AFTMC model for analysis, however, which may need a sufficient sample size.

Keywords: Cox's PH model, PHMC model, AFT model, AFTMC model, cure model

Background

The Cox's PH and AFT model are two most commonly used models for analyzing survival data[1, 2]. The underlying assumption of these two models is that each patient will eventually experience the event of interest (death, relapse or recurrence of cancer) if the follow-up time is sufficiently long[3-7]. However, with the development of medical science, advanced treatments can substantially improve survival rates for many diseases that are previously poor prognosis and even fatal, which means that a substantial proportion of subjects will never experience the event of interest[4-6]. An estimated Kaplan–Meier survival curve that presents a steady and long plateau at the tail and levels off at a value greater than zero with heavy censoring can be used as empirical evidence for the existence of cured individuals[3, 5, 8]. In fact, for most diseases, we are more concerned with the improvement of the cured rate rather than prolongation of time to the event of interest. For such time-to-event data, the cured rate can't be estimated at traditional methods such as Cox's PH or AFT model. Even some inappropriate conclusions may be drawn.

The cure model, which has been proposed by Boag (1949) and Berkson (1952)[9, 10] can be used to analyze such data. The mixture cure model assumes that the study population is a mixture of susceptible subjects, that are uncured or will experience the event of interest, and non-susceptible subjects, that are cured or will never experience the event of interest[11, 12]. At present, parametric approaches have been proposed for the mixture cure model by many scholars[12-14]. Since it is usually difficult to testify a parametric assumption, there has been increasing enthusiasm for the semiparametric mixture cure model[3, 5, 7, 8, 15]. An attractive characteristic of the mixture model is that it includes two parts which can be interpreted separately, in particular, the estimation of the proportion of the cured patients and of the failure time distribution of uncured patients.

However, for the data with a cured fraction, a great deal of literatures indicated that non-statistical professional researchers would still choose the traditional Cox's PH model or AFT model for

analysis. In view of this situation, this study intended to compare the performances of Cox's PH model with PHMC model and AFT model with AFTMC model respectively. In the meantime, two real data were analyzed to illustrate that for survival data with a cured fraction, the mixture cure model is a more effective method compared with traditional survival analysis methods.

Methods

The mixture cure model

Let Y be the indicator variable, where an individual will eventually ($Y = 1$) or never ($Y = 0$) experience the event of interest. When $Y = 1$, let T denote the time to the occurrence of an event of interest and $S(t|Y = 1)$ be the survival function of T . The mixture cure model can be denoted as:

$$S(t) = (1 - p) + p \times S(t | Y = 1) \quad (1)$$

Note that $S(t) \rightarrow 1 - p$ if $t \rightarrow \infty$. $p = Pr(Y = 1)$ represents the probability of a subject being uncured and is also referred to as "incidence". $S(t|Y = 1)$ is referred to as "latency". Assume that a logistic regression is employed to estimate the cured rate $1 - p(x)$

$$1 - p(x) = 1 - \frac{e^{(b_0 + b^T x)}}{1 + e^{(b_0 + b^T x)}} \quad (2)$$

Where b_0 is a constant, $b = (b_1, \dots, b_m)^T$ is a vector of unknown parameters and $x = (x_1, \dots, x_m)^T$ is a vector of covariates. If Cox's PH model is considered as "latency", which is

$$S(t | Y = 1, z) = S_0(t | Y = 1)^{e^{\beta^T z}} \quad (3)$$

the mixture cure model is called PHMC model. If AFT model is considered as "latency", which is

$$S(t | Y = 1, z) = S_0(t \times e^{\beta^T z}) \quad (4)$$

the mixture cure model is called AFTMC model. The Semiparametric Estimation Method of PHMC and AFTMC model can be seen the literatures[3, 5, 15, 16]. R package "smcure" was used to fit PHMC and AFTMC model.

Simulation Settings of Cox's PH and PHMC model

We generate the data according to the structure of PHMC model. The data should satisfy proportional hazards assumption, and the methods of tests and diagnostics are based on weighted residuals[17]. The incidence part of PHMC model is assumed that conditional probability of being

cured follows a logistic model:

$$1 - p(x) = 1 - \frac{e^{(b_0 + b_1 x_1 + b_2 x_2)}}{1 + e^{(b_0 + b_1 x_1 + b_2 x_2)}} \quad (5)$$

where $x_1 \sim N(0,1)$ is sampled from a standard normal distribution and $x_2 \sim Bernoulli(0.5)$ follow the binomial distribution. We assume $b_1 = b_2 = \log(2)$ and b_0 is selected such that the conditional probability of being cured are: 0 to 0.5 by 0.1. In the latency part of PHMC model, we generate $z_1 \sim N(0,1)$ that follows a standard normal distribution and $z_2 \sim Bernoulli(0.5)$ that obey the binomial distribution. Simultaneously, we consider $\beta_1 = \beta_2 = \log(2)$ corresponding to z_1 and z_2 respectively. According to the method proposed by Bender[18], we generate the survival time. We assume the baseline hazard function is from the Weibull distribution and both the scale and shape parameters are set to be 1. The censored time is from the exponential distribution, the parameter λ controls the censored rate, and the value of λ is adjusted so that the difference between the average censored rate and cured rate for each set of data are: 0.1 to 0.3 by 0.1. The sample size is set to be 100 and 200, and the number of simulations is 1000. The evaluation criteria include estimated biases, mean square error (MSE), confidence interval capture rate, and concordance probability K index[19, 20]. The estimated biases ($Bias = \bar{\beta} - \beta$) is the difference between the average estimated value of the model coefficient and the true value, which is used to evaluate the deviation of the coefficient estimation of model. The mean square error ($MSE = (\bar{\beta} - \beta)^2 + (SE(\hat{\beta}))^2$) is employed to evaluate the accuracy of the coefficient estimation of model. The confidence interval capture rate ($\hat{\beta}_i \pm Z_{1-\alpha/2} SE(\hat{\beta}_i)$) is the ratio of the confidence interval obtained to contain the true value of the coefficient of model, reflecting whether the model can control type I error. K index = $\text{pr}(T_1 > T_2 | R_2 > R_1 | Y_1 = Y_2 = 1)$ is utilized for measuring the discriminatory accuracy of the model.

Simulation Settings of the AFT and AFTMC model

We generate the data according to the structure of AFTMC model. In the incidence part of AFTMC model, we create two covariates x_1 and x_2 in the same way as simulation settings of Cox's PH and PHMC model and set coefficients of covariate $b_1 = b_2 = 2$. b_0 is selected such that the conditional probability of being cured are: 0 to 0.5 by 0.1. In the latency part of AFTMC model, we also generate two covariates z_1 and z_2 in the same manner as described above and set coefficients

of covariate $\beta_1 = \beta_2 = 2$. We generate the survival time based on the inverse function of Weibull distribution[21] and set scale parameter equal to be 1. The censored time is sampled from the exponential distribution. The rest part about simulation studies are all the same as simulation settings of Cox's PH and PHMC model.

Results

Simulation Results of Cox's PH and PHMC model

The estimated biases of two models are displayed in figure 1. When the cured rate is 0, that is, there is not cured individual, the absolute value of Bias in PHMC model are always close to Cox's PH model. When cured rates range of 0.1 to 0.5 by 0.1, the absolute values of Bias in PHMC model are always smaller than Cox's PH model. When the corresponding cured rate is fixed, with the increase of censored rate, the absolute value of Bias in both model gradually decreases, and the Cox's PH model changes more. It is suggested that the smaller the cured proportion in the censored individuals is. The smaller estimation bias of the coefficients will be.

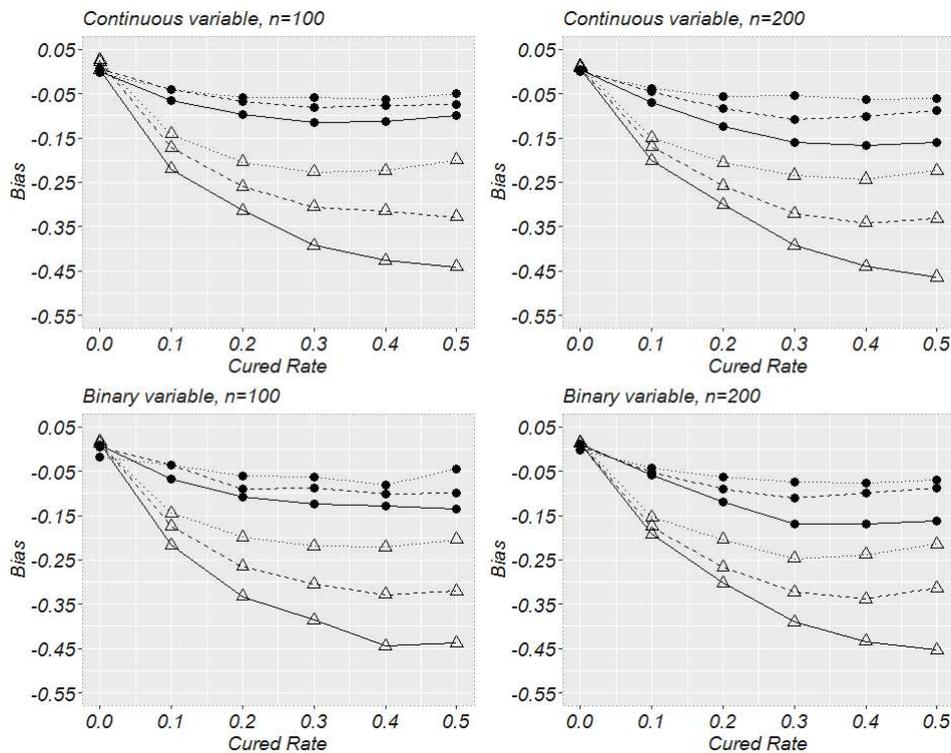


Figure 1. Estimated biases for Cox PH and PHMC model. The bias of PHMC model is its latency part. The triangles represent estimated biases of Cox PH model and the circles denote estimated biases of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

The MSE of two models are illustrated in figure 2. When the cured rate is 0, MSE is similar in

two models. With the sample size being 100, when both the cured and the censored rate are high (e.g. cured rate=0.5 and censored rate=0.8), MSE of PHMC model will become remarkably large, especially for binary covariate. In this case, PHMC model runs out of its accuracy. And when the sample size is 200, MSE of PHMC model decreases significantly, especially for continuous covariate, implying that PHMC model may require larger sample size than Cox's PH model. Except for the sample size of 100, high cured and censored rate, MSE of both models have a tendency to increase as the cured rate increases (0.1 to 0.5), and MSE of PHMC model is always smaller than Cox's PH model. When the corresponding cured rate is fixed, MSE of PHMC model is always smaller than Cox's PH model with the increase of censored rate. However, MSE of PHMC model is progressively increasing, while MSE of Cox's PH model is decreasing, which implies that the smaller the cured proportion in the censored part is, the smaller the advantage of PHMC model over Cox's PH model.

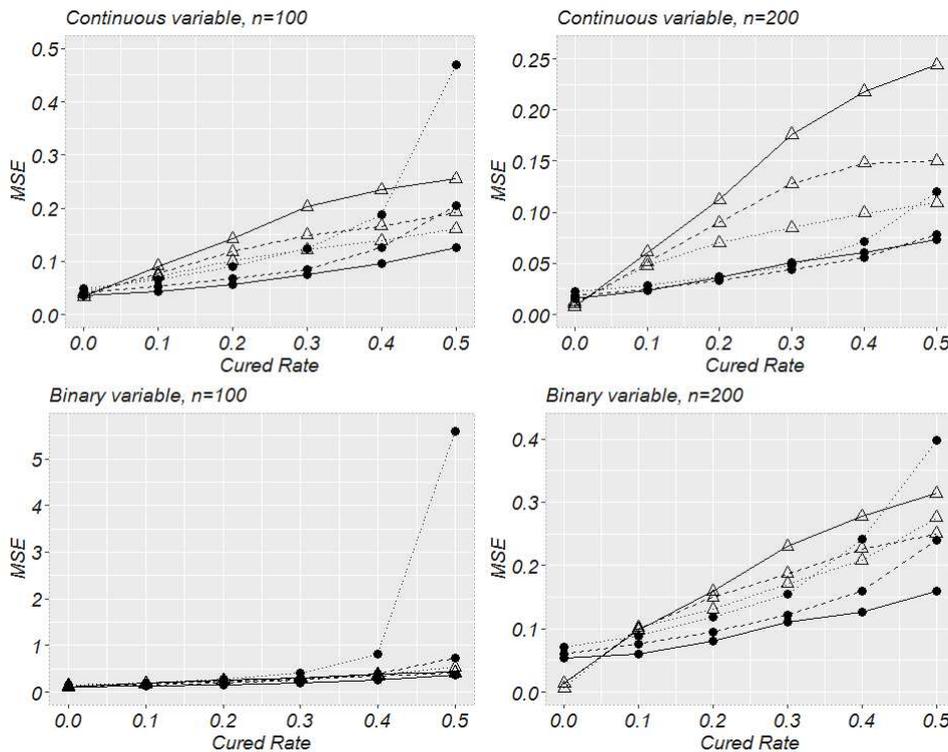


Figure 2. MSE for Cox PH and PHMC model. MSE of PHMC model is its latency part. The triangles represent MSE of Cox PH model and the circles denote MSE of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

The confidence interval capture rate (CP) of two models are exhibited in figure 3. The acceptable range of confidence interval capture rate is: $[p - 2SE(p), p + 2SE(p)]$, where $SE(p) = \sqrt{p(1-p)/B}$, p is capture rate and usually set to 0.95, B is the number of replications in

simulation study[22, 23]. Therefore, confidence interval capture rate is acceptable within the range of [0.936, 0.964] if the simulation study is carried out 1000 times. When the cured rate is 0, confidence interval capture rate for two models is mostly within the range of [0.936, 0.964], indicating that both models can control type I error well in this case. The confidence interval capture rate of PHMC model varies little under various circumstances, which is always close to the acceptable range. With the increase of cured rate, confidence interval capture rate of Cox's PH model shows a tendency of first significantly decreasing, then stable or rise again. When the corresponding cured rate is fixed, confidence interval capture rate of Cox's PH model becomes closer to the acceptable range as the censored rate increases.

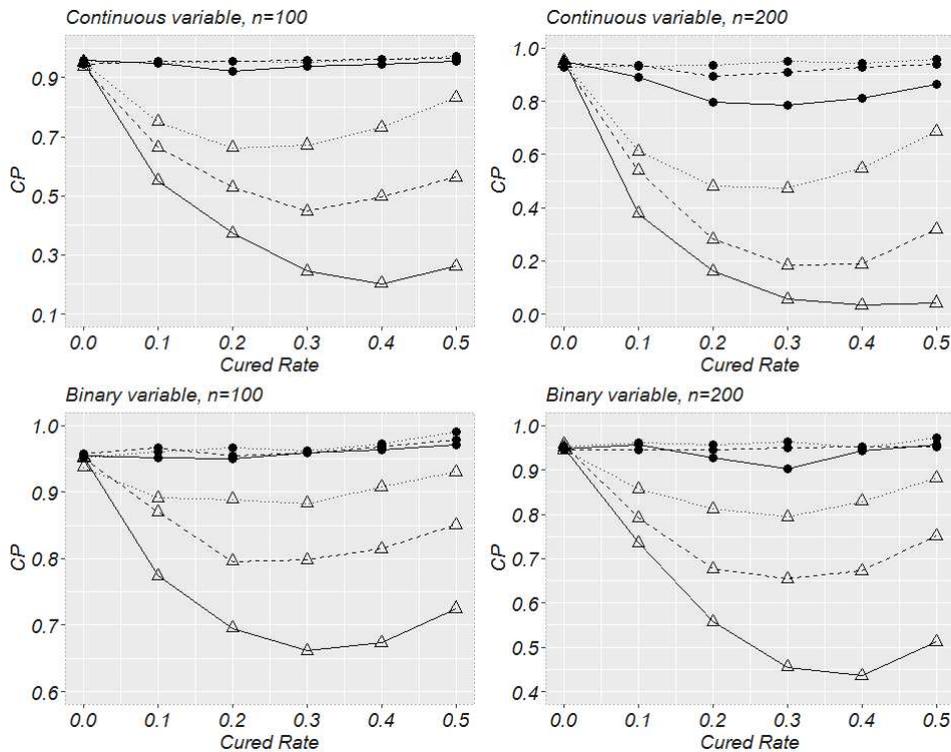


Figure 3. Confidence interval capture rate (CP) for Cox PH and PHMC model. Confidence interval capture rate of PHMC model is its latency part. The triangles represent confidence interval capture rate of Cox PH model and the circles denote confidence interval capture rate of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

The K indexes of two models are presented in figure 4. When the cured rate is 0, the K indexes of two models are almost equal. When the cured rate ranges from 0.1 to 0.5, the K indexes of PHMC model are always greater than Cox's PH model, indicating that when survival data has a substantial proportion of subjects being cured, the overall fit effect of PHMC model is better. When the corresponding cured rate is fixed, the K indexes of PHMC model are greater than that of Cox's PH

model. With the increase of censored rate, the K indexes of both PHMC and Cox's PH model will increase, implying that the smaller the cured proportion in censored individuals is, the better the model fit effect are both in the two models.

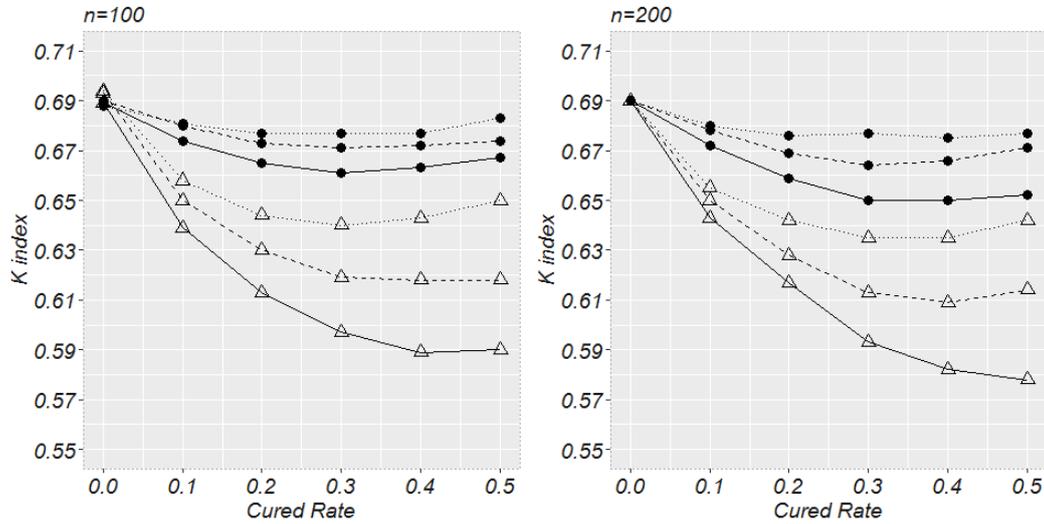


Figure 4. K index for Cox PH and PHMC model. The triangles represent the K index of Cox PH model and the circles denote the K index of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

In summary, when the cured rate is 0, the estimated bias, confidence interval capture rate, and K index of PHMC model is close to Cox's PH model, but MSE of PHMC model is slightly larger than Cox's PH model. Therefore, in this case, PHMC model has no obvious advantage. When survival data has a substantial proportion of subjects being cured, the absolute value of Bias and MSE in PHMC model are always smaller than Cox's PH model, confidence interval capture rate of PHMC model is always closer to the acceptable range than Cox's PH model, and the K index of PHMC model is always greater than Cox's PH model. Consequently, in this situation, it is recommended to utilize PHMC model for analysis, but it may need a sufficient sample size.

Simulation Results of the AFT and AFTMC model

The estimated biases of two models are exhibited in figure 5. When the cured rate is 0, estimated biases of AFTMC and AFT model are still close to zero. In all other situations, estimated biases of AFTMC model are always around zero. As the cured rate increases, the absolute value of Bias in the AFT model increases rapidly, and then gets stable or decreases. When the corresponding cured rate is fixed, the absolute value of Bias in the AFT model decreases as the censored rate increases.

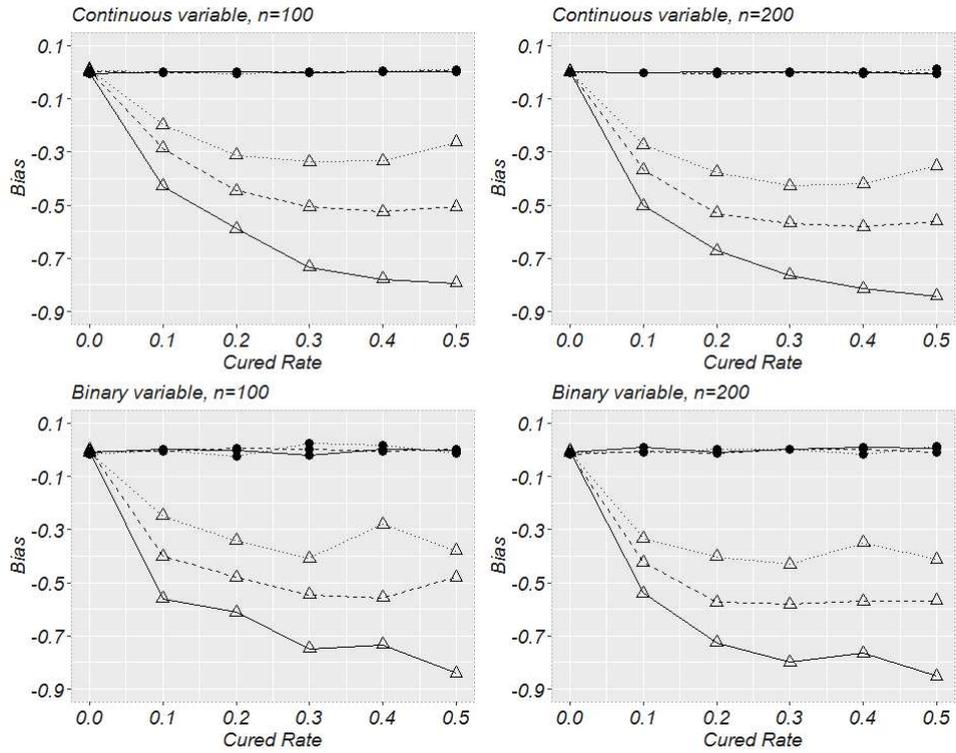


Figure 5. Estimated biases for AFT and AFTMC model. Bias of AFTMC model is its latency part. The triangles represent estimated biases of AFT model and the circles denote estimated biases of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

The MSE of two models are exhibited in figure 6. When the cured rate is 0, MSE of two models are close to 0. When the cured rate is from 0.1 to 0.5, MSE of two models are gradually increasing, MSE of AFTMC model is relatively stable, and MSE of AFT model has a wide range of changes. Simultaneously, MSE of AFTMC model is always smaller than AFT model. When the corresponding cured rate is fixed, MSE of AFTMC model increases as the censored rate increases, but MSE of AFT model decreases.

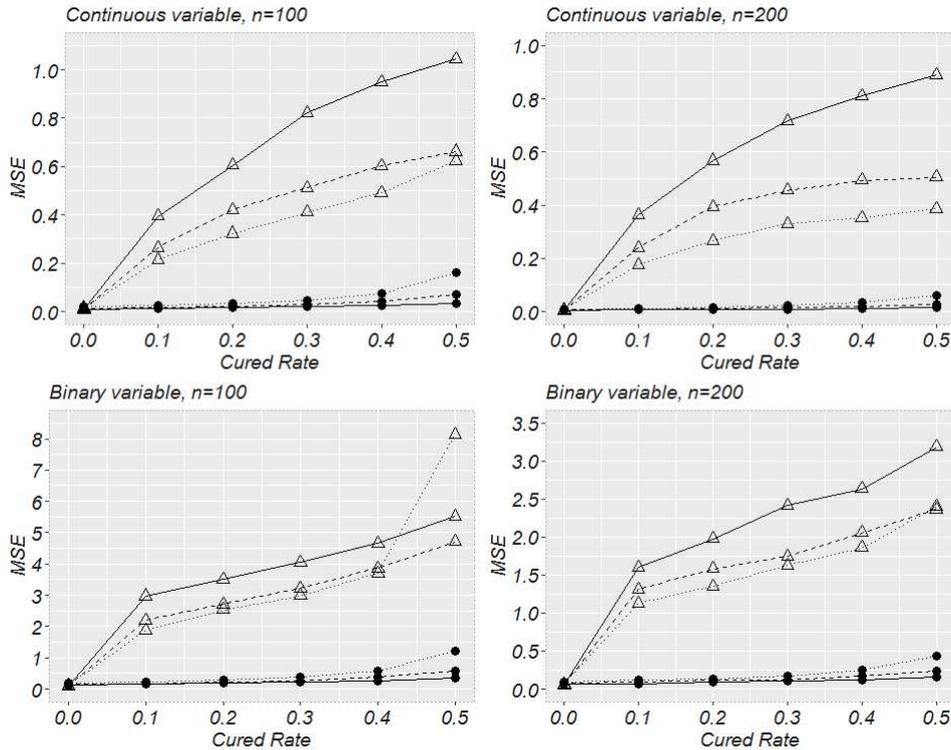


Figure 6. MSE for AFT and AFTMC model. MSE of AFTMC model is its latency part. The triangles represent MSE of AFT model and the circles denote MSE of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

The confidence interval capture rate of two models are displayed in figure 7. When the cured rate is 0, confidence interval capture rate for two models are mostly within the range of [0.936, 0.964], indicating that both models can control type I error well in this case. The confidence interval capture rate of PHMC model varies little under various circumstances and is always close to the acceptable range. When the cured rate is 0.1 to 0.5, confidence interval capture rate of AFT model is low. With the increase of the cured rate, confidence interval capture rate of the binary variables of AFT model is getting closer to the acceptable range but for continuous variables, the change of confidence interval capture rate is very small. When the corresponding cured rate is fixed, confidence interval capture rate gradually increases as the censored rate increases.

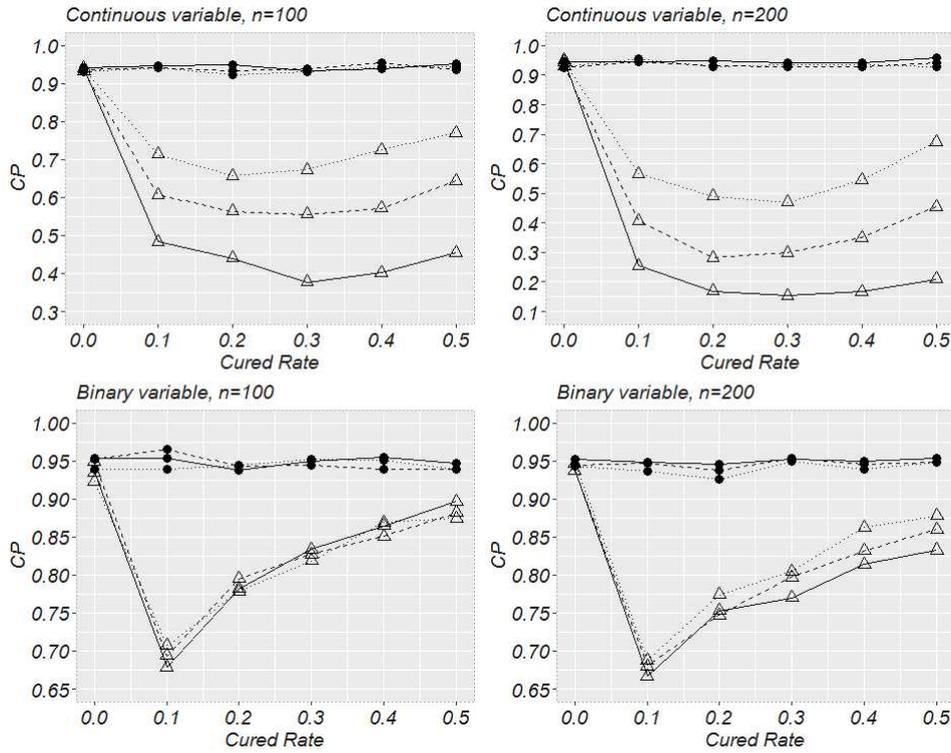


Figure 7. Confidence interval capture rate (CP) for AFT and AFTMC model. Confidence interval capture rate of AFTMC model is its latency part. The triangles represent confidence interval capture rate of AFT model and the circles denote confidence interval capture rate of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

The K indexes of two models are shown in figure 8. When the cured rate is 0, the K indexes of two models are similar. In all cases, the K indexes of AFTMC model are consistently stable at a high level (0.91). With the increase of the cured rate, the K index of AFT model gradually decreases, indicating that the model fit effect of AFTMC model is better than AFT model if there are cured individuals. When the corresponding cured rate is fixed, the K index of AFT model gradually increases as the censored rate increases.

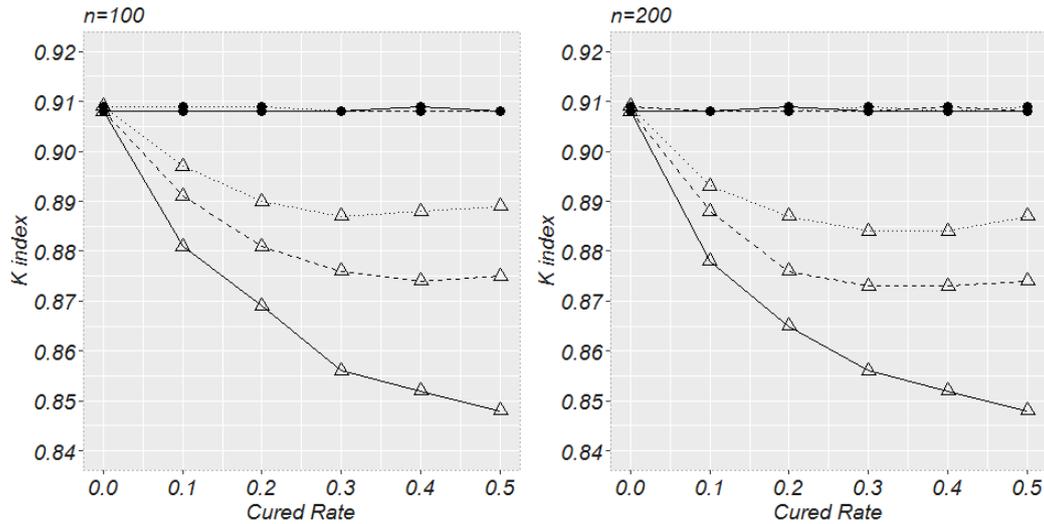


Figure 8. K index for AFT and AFTMC model. The triangles represent the K index of AFT model and the circles denote the K index of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

In summary, when the cured rate is 0, Estimated bias, MSE, confidence interval capture rate, and K index of AFTMC model are similar to AFT model. Therefore, AFTMC model does not possess a clear advantage in this case. When survival data has a substantial proportion of subjects being cured, the absolute value of Bias and MSE of AFTMC model are always smaller than AFT model, confidence interval capture rate of AFTMC model is always closer to the acceptable range than AFT model, and K index of AFTMC model is always larger than AFT model. Consequently, the AFTMC model is recommended for analysis if there are cured individuals in survival data.

Examples

Lung Cancer Data

Survival of patients with lung cancer is from the North Central Cancer Treatment Group[24]. This study was intended to explore factors that may affect patients' prognosis. The data contained 228 patients. Covariates included Survival time, censored status, age, sex, ECOG score, Karnofsky performance score rated by doctor, Karnofsky performance score rated by patient, calorie consumed in meals, and weight loss in the past six months. K-M survival curves of male and female patients are presented in figure 5, which shows that after approximately 1000 days of follow-up, the survival rate is almost 0, implying that there is not cured individual.

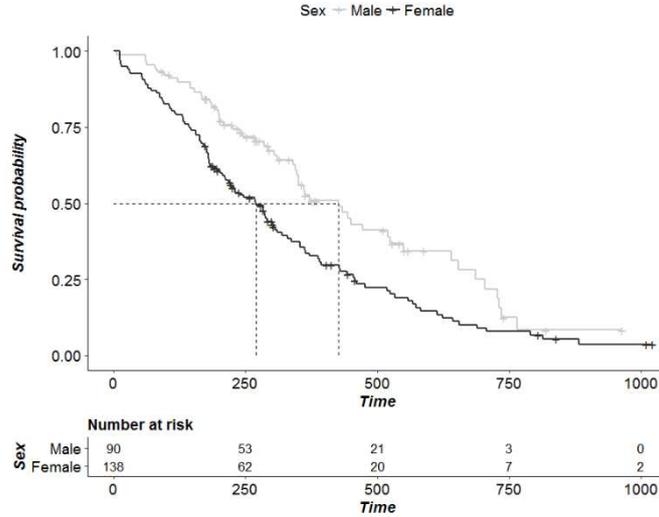


Figure 9. K-M survival curves for lung cancer patients. Time(days)

With LASSO regression[25, 26], the covariates selected were sex, ECOG performance score, and Karnofsky performance score rated by patients. The proportional hazards test exhibits the data complies with the proportional hazards assumption ($P=0.165$). We used Cox's PH and PHMC model to analyze it respectively. The results are exhibited in Table 1. The results of the incidence part of PHMC model show that three covariates are not statistically significant (P values were greater than 0.9) and aren't displayed in Table1. As showed from Table 1, both models display that sex is statistically significant, and pat.karno (Karnofsky performance score rated by patient) is not statistically significant. Though ph.ecog (ECOG performance score) displays statistical significance in Cox's PH model and no statistical significance in PHMC model, the P value is close to the significant level of 0.05. It can be considered that the analysis results of two models are similar.

Table1. The result for Lung Cancer Data

Effect	PHMC Model(Latency)			Cox Model		
	$\hat{\beta}$	HR (95%CI)	P value	$\hat{\beta}$	HR (95%CI)	P value
sex	-0.511	0.600(0.401,0.897)	0.013	-0.540	0.583(0.418,0.812)	0.001
ph.ecog	0.313	1.367(0.981,1.906)	0.065	0.389	1.476(1.128,1.931)	0.005
pat.karno	-0.008	0.992(0.975,1.009)	0.345	-0.009	0.991(0.978,1.005)	0.198

ph.ecog: ECOG performance score; pat.karno: Karnofsky performance score rated by patient

Eastern Cooperative Oncology Group (ECOG) Data

Melanoma data are taken from the Eastern Cooperative Oncology Group (ECOG) phase III clinical trial E1684[27]. The purpose of the E1684 clinical trial was to evaluate the high dose interferon alpha-2b (IFN) regimen (treatment group) against the placebo (control group) as the postoperative

adjuvant therapy on relapse-free survival in patients with American Joint Committee on Cancer (AJCC) stage IIB or III melanoma. E1684 contained 284 patients with no missing value and included covariates such as observed relapse-free time, censored status, AGE, SEX, and TRT (0=control group, 1=treatment group). The K-M survival curve of the treatment group and the control group is presented in figure 6, which show a steady and long plateau at the tail, implying that there may be cured individuals in melanoma patients.

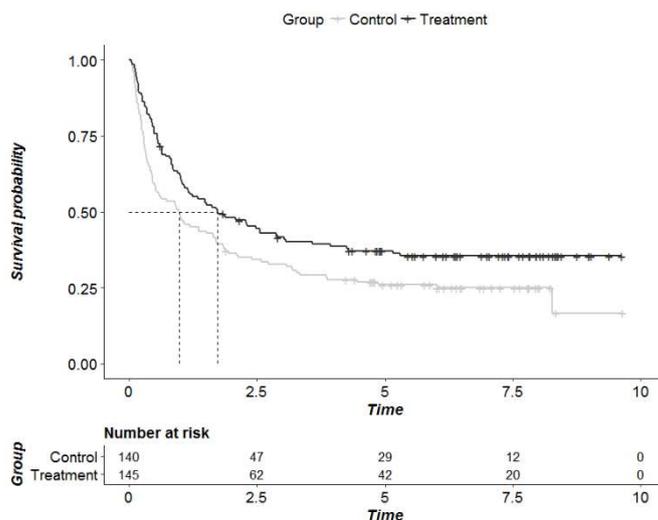


Figure 10. K-M survival curves for E1684 patients. Time(years)

Proportional hazards test shows the proportional hazards assumption is met ($P=0.404$). Then Cox’s PH and PHMC model were used for analysis. The results are presented in table 2 and table 3. In table 2, the results of the incidence part in PHMC model exhibit that TRT ($OR=0.555$, $P=0.053$) may be a factor that affects the cured rate of melanoma patients, suggesting that high dose interferon alpha-2b (IFN) may increase the cured rate of melanoma patients.

Both models in Table 3 present that SEX and AGE are not statistically significant. For TRT, no statistical significance is exhibited in the latency part of PHMC model ($HR=0.858$, $P =0.313$), indicating that high dose interferon alpha-2b (IFN) doesn’t delay the occurrence of recurrent events compared with placebo or can’t reduce the risk of recent recurrence. However, in Cox’s PH model, TRT is statistically significant and its HR is $0.698(95\%CI: 0.526\sim0.924, P=0.012)$. This result may be summarized either as “treatment group therapy is associated with a 30% ($= 1-0.70$) reduction in the risk of relapse”, or as “treatment group therapy prevents 30% of the relapse that would occur without the therapy”. Such statements are generally found in leading medical journals. Hence, given this information, most patients would believe their chance of being cured could be increased with

the treatment if it shows a *HR* less than 1. However, $1-HR$ is not the proportion of patients whose relapse are prevented by treatment, but the proportion of relapse that are either prevented or delayed. Thus, the *HR* in Cox's PH model cannot distinguish between treatments that prevent relapse (curative treatment) and those that merely delay its relapse.

Table2. The result for PHMC Model (Incidence)

Effect	PHMC Model (Incidence)		
	$\hat{\beta}$	OR (95%CI)	P value
Intercept	1.365	3.915(2.320,6.607)	<0.001
TRT	-0.588	0.555(0.306,1.007)	0.053
SEX	-0.087	0.917(0.432,1.944)	0.821
AGE	0.020	1.021(0.992,1.050)	0.158

TRT: (0=control group, 1=treatment group)

Table3. The result for E1684 patients

Effect	PHMC Model (Latency)			Cox Model		
	$\hat{\beta}$	HR (95%CI)	P value	$\hat{\beta}$	HR (95%CI)	P value
TRT	-0.154	0.858(0.637,1.155)	0.313	-0.360	0.698(0.526,0.924)	0.012
SEX	0.099	1.105(0.772,1.581)	0.587	-0.018	0.982(0.737,1.310)	0.903
AGE	-0.008	0.992(0.979,1.005)	0.250	0.005	1.005(0.995,1.015)	0.357

TRT: (0=control group, 1=treatment group)

Discussion

Although the cure model has been proposed for many years and the theory is relatively perfect, it hasn't been widely used in the clinical literature. For diseases with long-term survival or even cure, the cure model is a very effective method for analysis. We carry out simulation studies to compare the performances of Cox's PH model with PHMC model and AFT model with AFTMC model respectively in this article. The results show that PHMC and AFTMC model do not have obvious advantages for analyzing survival data without a cured fraction. But when subjects exist in the data, who will never experience the event of interest, it is recommended to use PHMC or AFTMC model for analysis, which may need relatively larger sample size. Case studies display that when survival data have not cured individuals, the results of two models are similar. However, the conclusions of two models may be different for analyzing data with a substantial proportion of cured subjects. The mixture cure model can explore the effects of covariates both on delaying events and preventing events, so that a more specific and clear explanation can be obtained. The example analysis matches

the simulation results.

The results of this study and related literatures [Maetani S, Gamel J W. Parametric Cure Model versus Proportional Hazards Model in Survival Analysis of Breast Cancer and Other Malignancies[J]. *Advances in Breast Cancer Research*, 2013, 2(4):119-125.] suggest that traditional methods of survival analysis (Cox or AFT) fail to distinguish between curative and life-prolonging treatments. The methods of traditional survival analysis are more sensitive to find out treatments that delay the occurrence of events (higher power) but are less sensitive to treatments that increase the cured rate, especially when the length of follow-up is limited, giving doctors a false estimate of the best treatment. For those studies with long follow-up and the existence of cured individuals, the power of the traditional survival analysis methods will reduce. Therefore, unless all patients will eventually experience the event of interest or there are no long-term survivors, the primary measure to evaluate patient's benefit should be the patient's cured rate rather than the risk ratio. Compared with traditional Cox or AFT model, which only focuses on the time of event and analysis results may mislead doctors or patients, the cure model can provide more specific and accurate information about the long-term survival benefit of the treatment.

In this article, the difference between the average censored and average cured rate is set from 0.1 to 0.3, and scenarios of greater censored rate are not involved in this study. Additionally, appropriate sample size for cure model has not yet been studied. Furthermore, we just concentrate on mixture cure models in this article and non-mixture cure models are not involved in our comparison.

A competing risks derivation of a mixture cure model has been proposed by JB. GreenHouse as early as in 1984[28]. Semiparametric accelerated failure time cure rate mixture model with competing risks has been put forward recently[29]. Thus it is worth further research to compare the competing risks model with its mixture model.

Conclusion

The PHMC model and AFTMC model do not have obvious benefits for time-to-event data without a cured fraction. In this case, it is recommended to utilize Cox's PH or AFT model for analysis. If some subjects are non-susceptible to the event of interest in the data, it is recommended to utilize PHMC or AFTMC model for analysis, however, which may need a sufficient sample size.

Abbreviations

PHMC model: proportional hazards mixture cure model; AFTMC model: AFT mixture cure model;

MSE: mean square error

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analysed during this study are included in these published article.

1. Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group.
2. Interferon alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: the Eastern Cooperative Oncology Group Trial EST 1684.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was funded in part by Plan of Southern Medical University Clinical Research (LC2016YM016).

Authors' contributions

ASL and HFQ were the major contributor in design of simulation, writing code, and writing and revise the manuscript. KP and LYX interpreted the data and were the writing assistance. All authors have contributed to, read and approved the final manuscript for submission.

Acknowledgements

Not applicable.

References

1. Cox DR: **Regression Models and Life-Tables.** *Journal of the Royal Statistical Society* 1972, **34**(2):187-220.
2. Wei LJ: **The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis.** *STAT MED* 1992, **11**(14-15):1871-1879.
3. Sy JP: **Estimation in a Cox Proportional Hazards Cure Model.** *BIOMETRICS* 2000, **56**(1):227-236.
4. Peng Y, Zhang J: **Estimation method of the semiparametric mixture cure gamma frailty model.** *STAT MED* 2008, **27**(25):5177.
5. Zhang J, Peng Y: **A new estimation method for the semiparametric accelerated failure time mixture cure model.** *COMPUT STAT DATA AN* 2009, **38**(9):1980-1990.
6. Corbière F, Commenges D, Taylor JM, Joly P: **A penalized likelihood approach for mixture cure models.** *STAT MED* 2009, **28**(3):510.
7. Taylor JMG: **Semi-Parametric Estimation in Failure Time Mixture Models.** *BIOMETRICS* 1995, **51**(3):899-907.
8. Li CS, Taylor JMG: **A semi-parametric accelerated failure time cure model.** *STAT MED* 2002, **21**(21):3235.
9. Boag JW: **Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy.** *Journal of the Royal Statistical Society* 1949, **11**(1):15-53.
10. Joseph Berkson, Gage R: **Survival Curve for Cancer Patients Following Treatment.** *Publications of the American Statistical Association* 1952, **47**(259):501-515.
11. Corbière F, Joly P: **A SAS macro for parametric and semiparametric mixture cure models.** *COMPUT METH PROG BIO* 2007, **85**(2):173-180.
12. Farewell VT: **The use of mixture models for the analysis of survival data with long-term survivors.** *BIOMETRICS* 1982, **38**(4):1041.
13. Peng Y, Dear KBG, Denham JW: **A generalized F mixture model for cure rate estimation.** *STAT MED* 2010, **17**(8):813-830.
14. Yamaguchi K: **Accelerated Failure-Time Regression Models with a Regression Model of Surviving Fraction: An Application to the Analysis of "Permanent Employment" in Japan.** *Publications of the American Statistical Association* 1992, **87**(418):284-292.
15. Peng Y: **Fitting semiparametric cure models.** *COMPUT STAT DATA AN* 2003, **41**(3):481-490.
16. Peng Y: **A Nonparametric Mixture Model for Cure Rate Estimation.** *BIOMETRICS* 2000, **56**(1):237-243.
17. Grambsch PM, Therneau TM: **Proportional hazards tests and diagnostics based on weighted residuals.** *BIOMETRIKA* 1994, **81**(3):515-526.
18. Bender R, Augustin T, Blettner M: **Generating survival times to simulate Cox proportional hazards models.** *STAT MED* 2005, **24**(11):1713-1723.
19. Gönen M, Heller G: **Concordance Probability and Discriminatory Power in Proportional Hazards Regression.** *BIOMETRIKA* 2005, **92**(4):965-970.
20. Zhang Y, Shao Y: **Concordance measure and discriminatory accuracy in transformation cure models.** *BIostatistics* 2017.
21. Zhou X, Yan L, Prows DR, Yang R: **Generalized F accelerated failure time model for mapping survival trait loci.** *GENOMICS* 2011, **97**(6):379-385.
22. Tang L, Song J, Belin TR, Unützer J: **A comparison of imputation methods in a longitudinal**

- randomized clinical trial.** *STAT MED* 2005, **24**(14):2111-2128.
23. Burton A, Altman DG, Royston P, Holder RL: **The design of simulation studies in medical statistics.** *STAT MED* 2006, **25**(24):4279-4292.
24. Loprinzi CL, Laurie JA, Wieand HS, Krook JE, Novotny PJ, Kugler JW, Bartel J, Law M, Bateman M, Klatt NE: **Prospective evaluation of prognostic variables from patient-completed questionnaires. North Central Cancer Treatment Group.** *Journal of Clinical Oncology Official Journal of the American Society of Clinical Oncology* 1994, **12**(3):601-607.
25. Hossain, Shakhawat, Ahmed, Ejaz S: **Penalized and Shrinkage Estimation in the Cox Proportional Hazards Model.** *Communications in Statistics* 2014, **43**(5):1026-1040.
26. Tibshirani R: **Regression shrinkage and selection via the lasso: a retrospective.** *Journal of the Royal Statistical Society: Series B Statistical Methodology* 2011, **73**(3):273-282.
27. Kirkwood JM, Strawderman MH, Ernstoff MS, Smith TJ, Borden EC, Blum RH: **Interferon alfa-2b adjuvant therapy of high-risk resected cutaneous melanoma: the Eastern Cooperative Oncology Group Trial EST 1684.** *J CLIN ONCOL* 1996, **14**(1):7-17.
28. Greenhouse JB, Wolfe RA: **A competing risks derivation of a mixture model for the analysis of survival data.** *Communications in Statistics* 1984, **13**(25):3133-3154.
29. Choi S, Zhu L, Huang X: **Semiparametric accelerated failure time cure rate mixture models with competing risks.** *STAT MED* 2017, **37**(1).

Figures

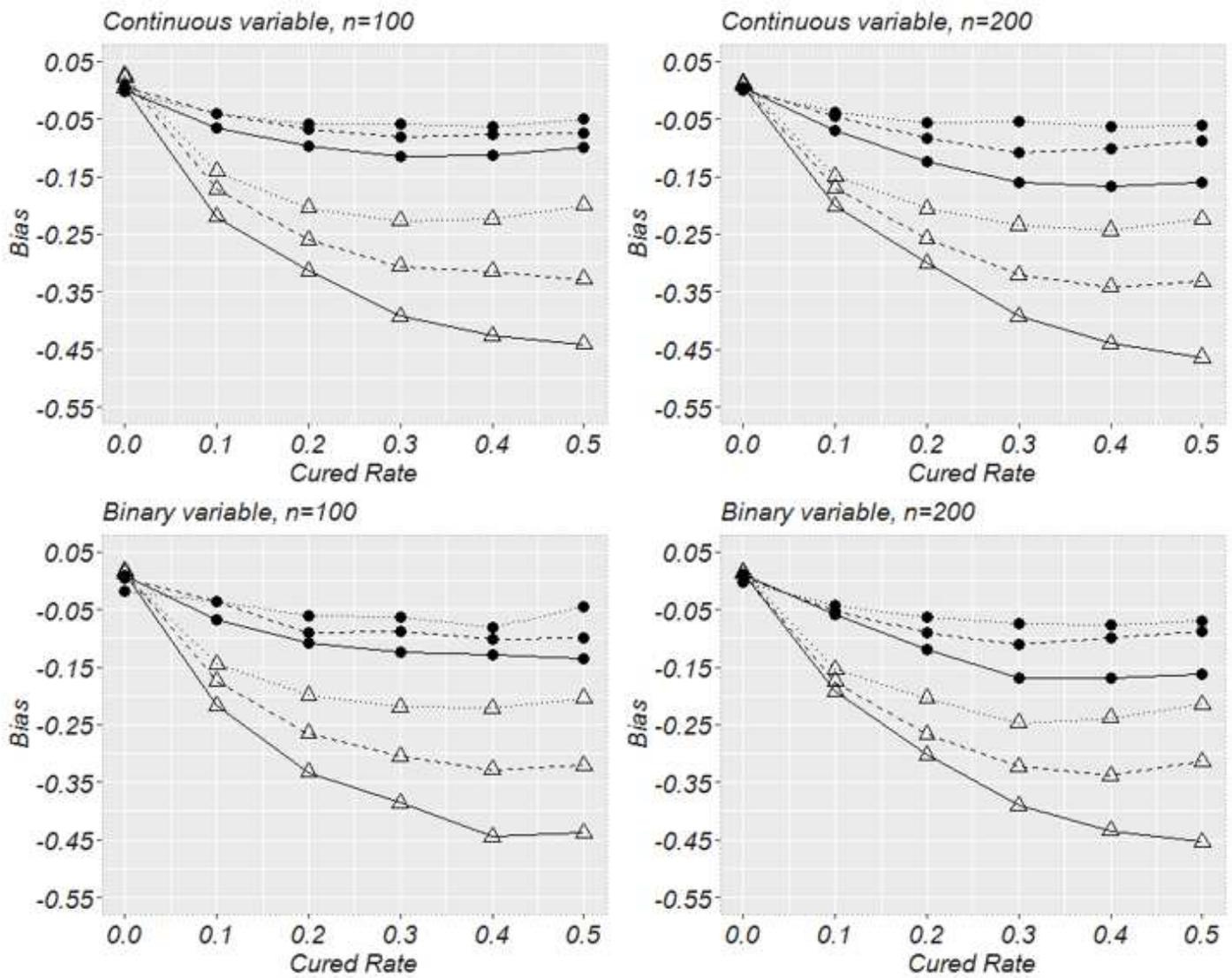


Figure 1

Estimated biases for Cox PH and PHMC model. The bias of PHMC model is its latency part. The triangles represent estimated biases of Cox PH model and the circles denote estimated biases of PHMC model. Solid, dashed, and dotted line represent censored rate-cured rate=0.1, 0.2 and 0.3, respectively.

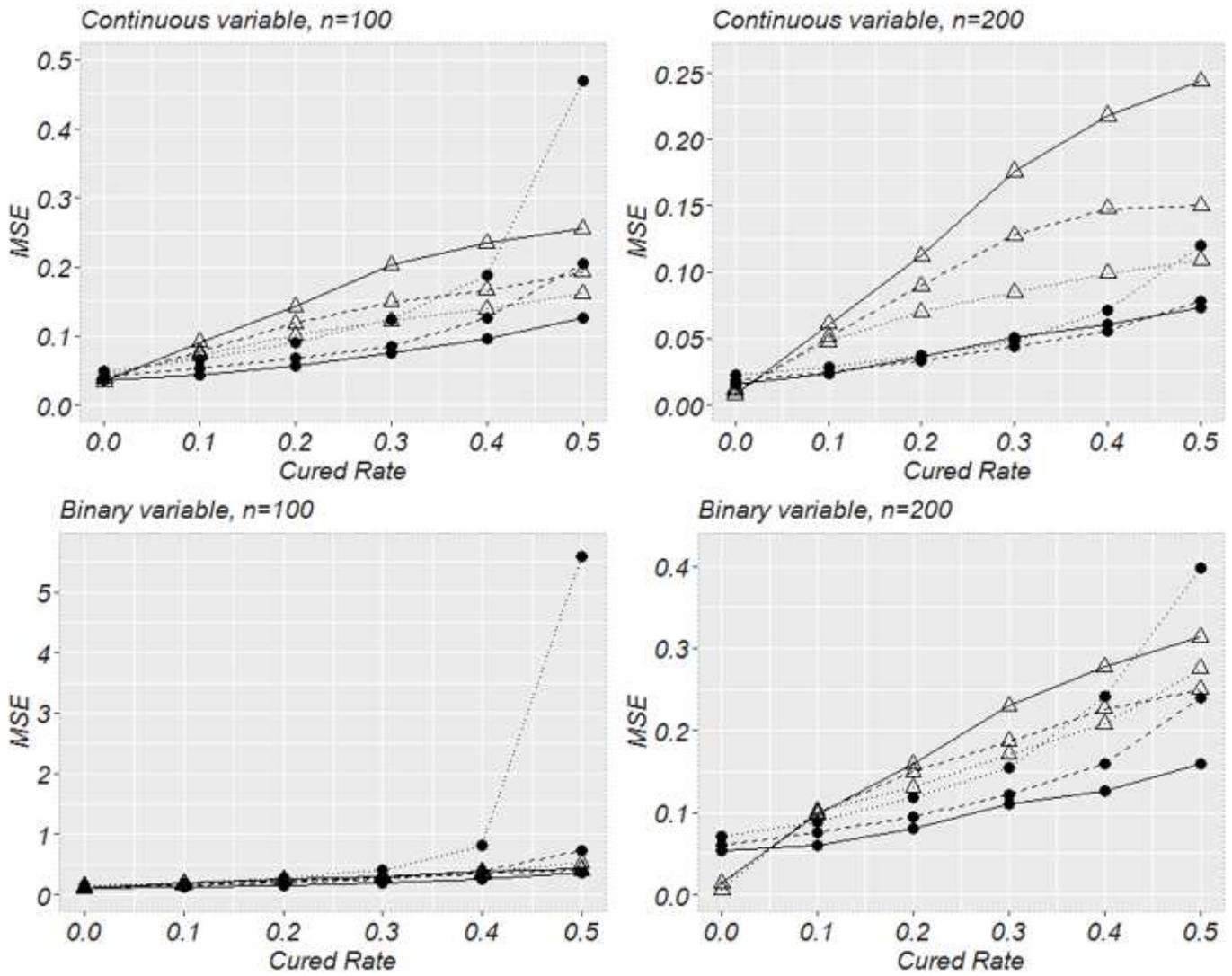


Figure 2

MSE for Cox PH and PHMC model. MSE of PHMC model is its latency part. The triangles represent MSE of Cox PH model and the circles denote MSE of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

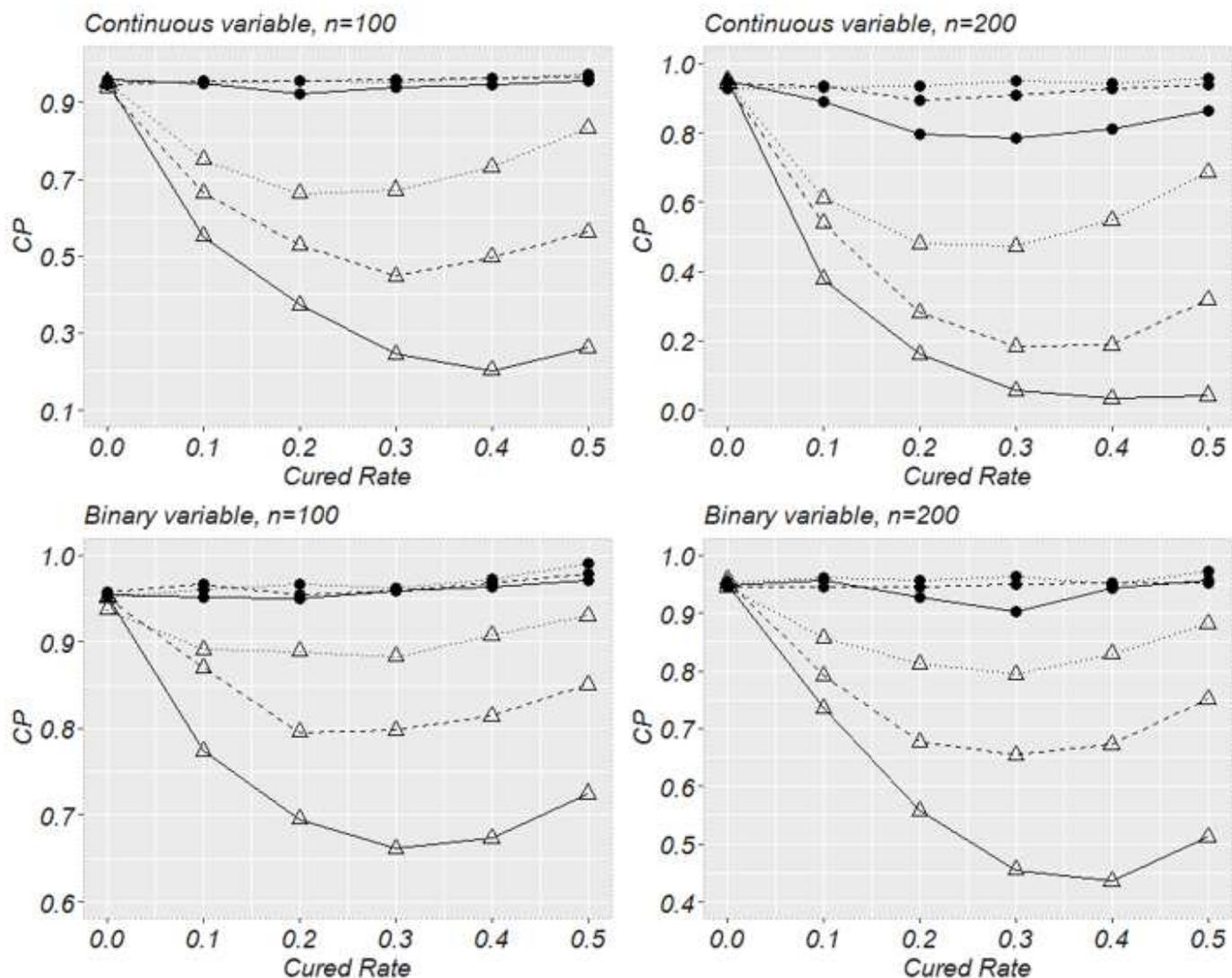


Figure 3

Confidence interval capture rate (CP) for Cox PH and PHMC model. Confidence interval capture rate of PHMC model is its latency part. The triangles represent confidence interval capture rate of Cox PH model and the circles denote confidence interval capture rate of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

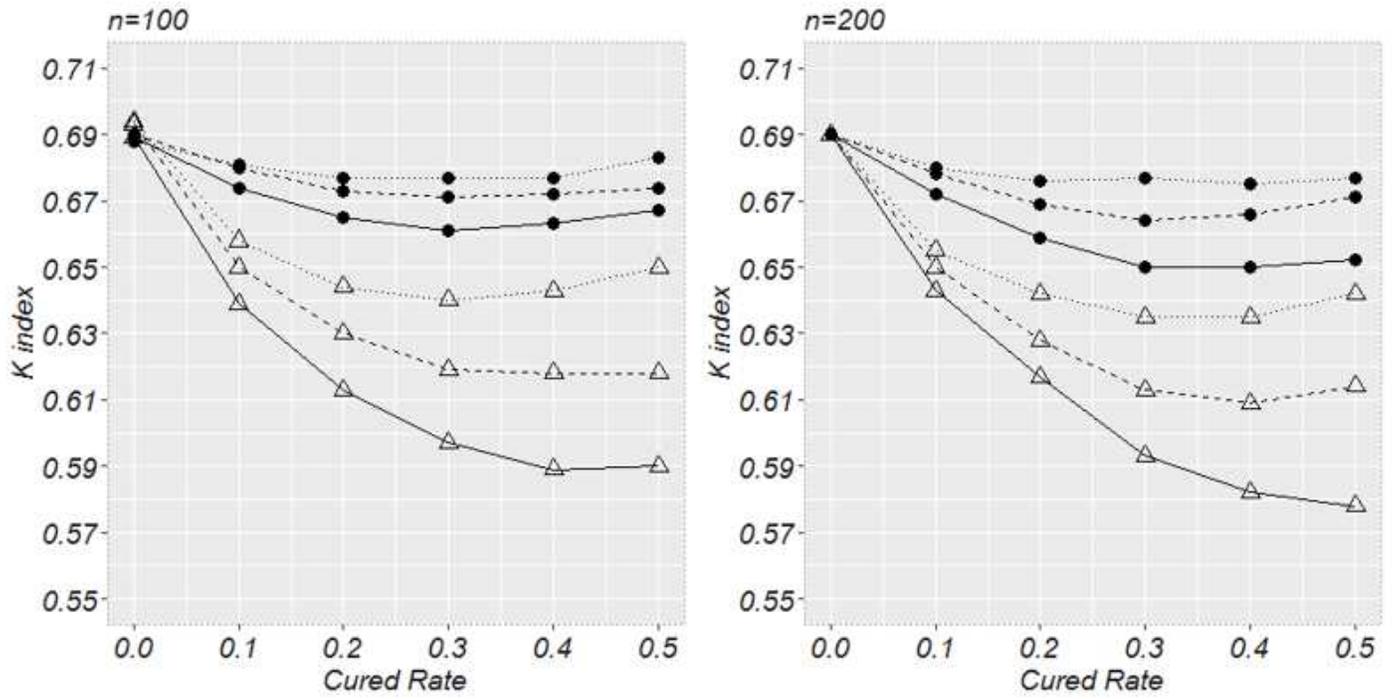


Figure 4

K index for Cox PH and PHMC model. The triangles represent the K index of Cox PH model and the circles denote the K index of PHMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

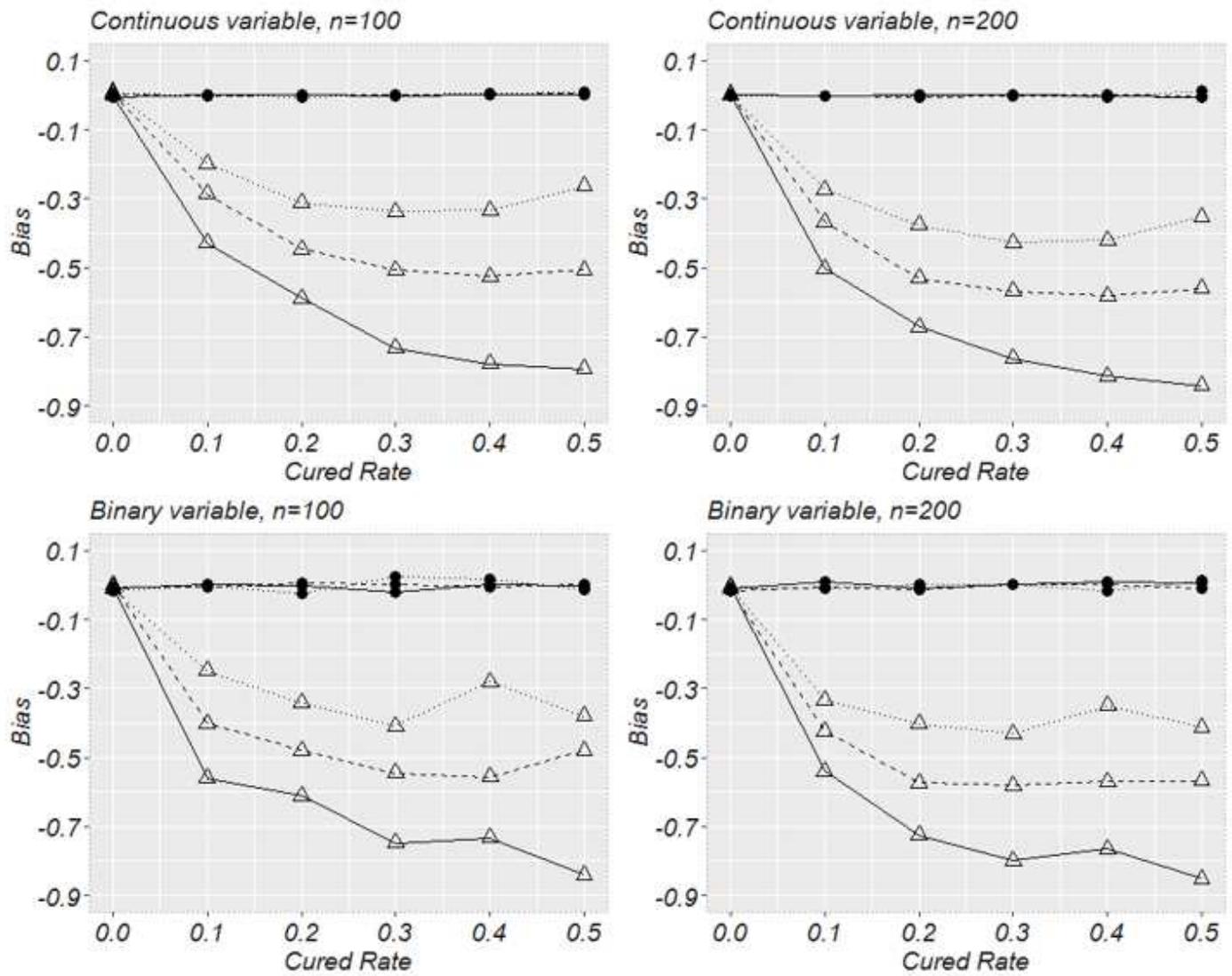


Figure 5

Estimated biases for AFT and AFTMC model. Bias of AFTMC model is its latency part. The triangles represent estimated biases of AFT model and the circles denote estimated biases of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

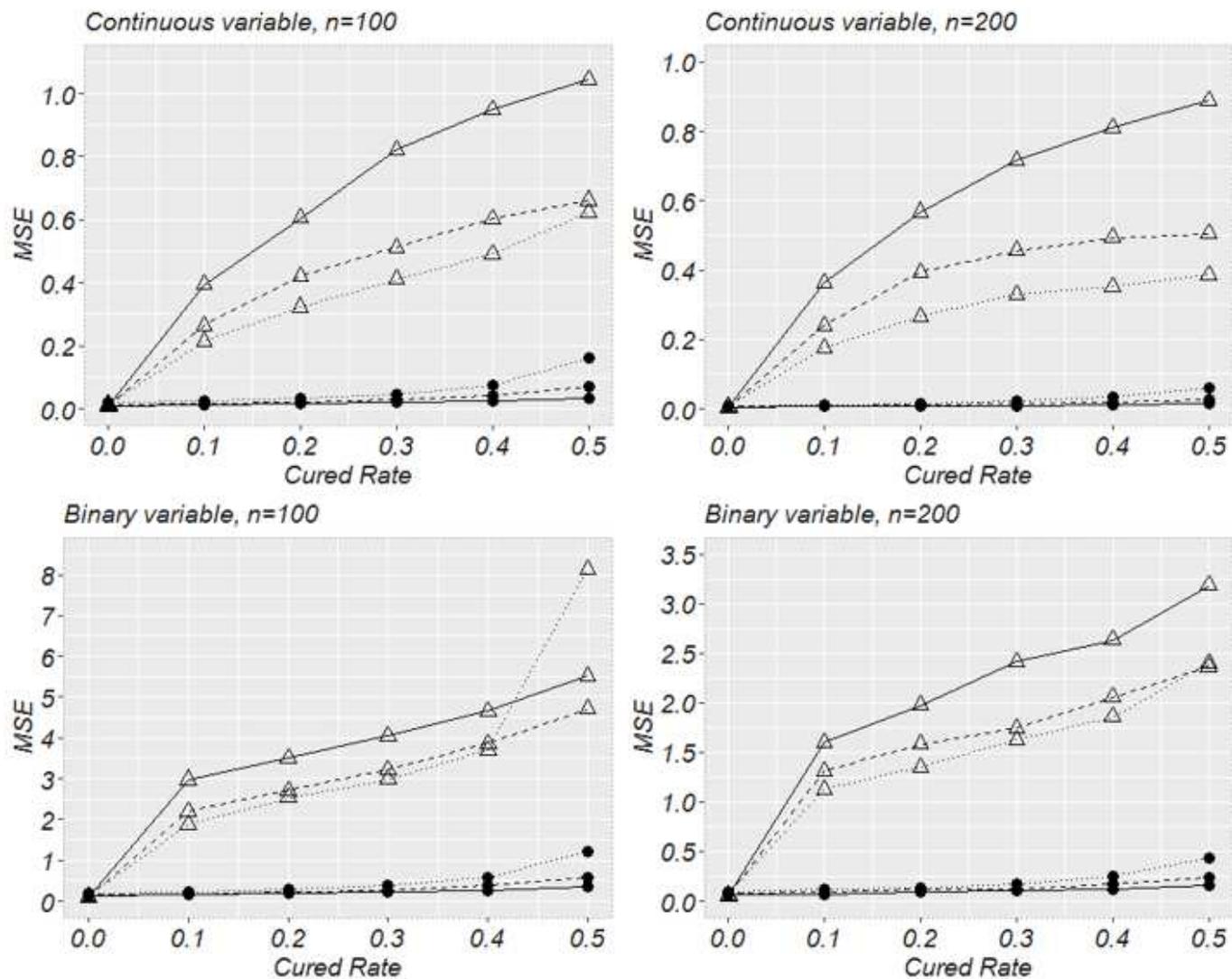


Figure 6

MSE for AFT and AFTMC model. MSE of AFTMC model is its latency part. The triangles represent MSE of AFT model and the circles denote MSE of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

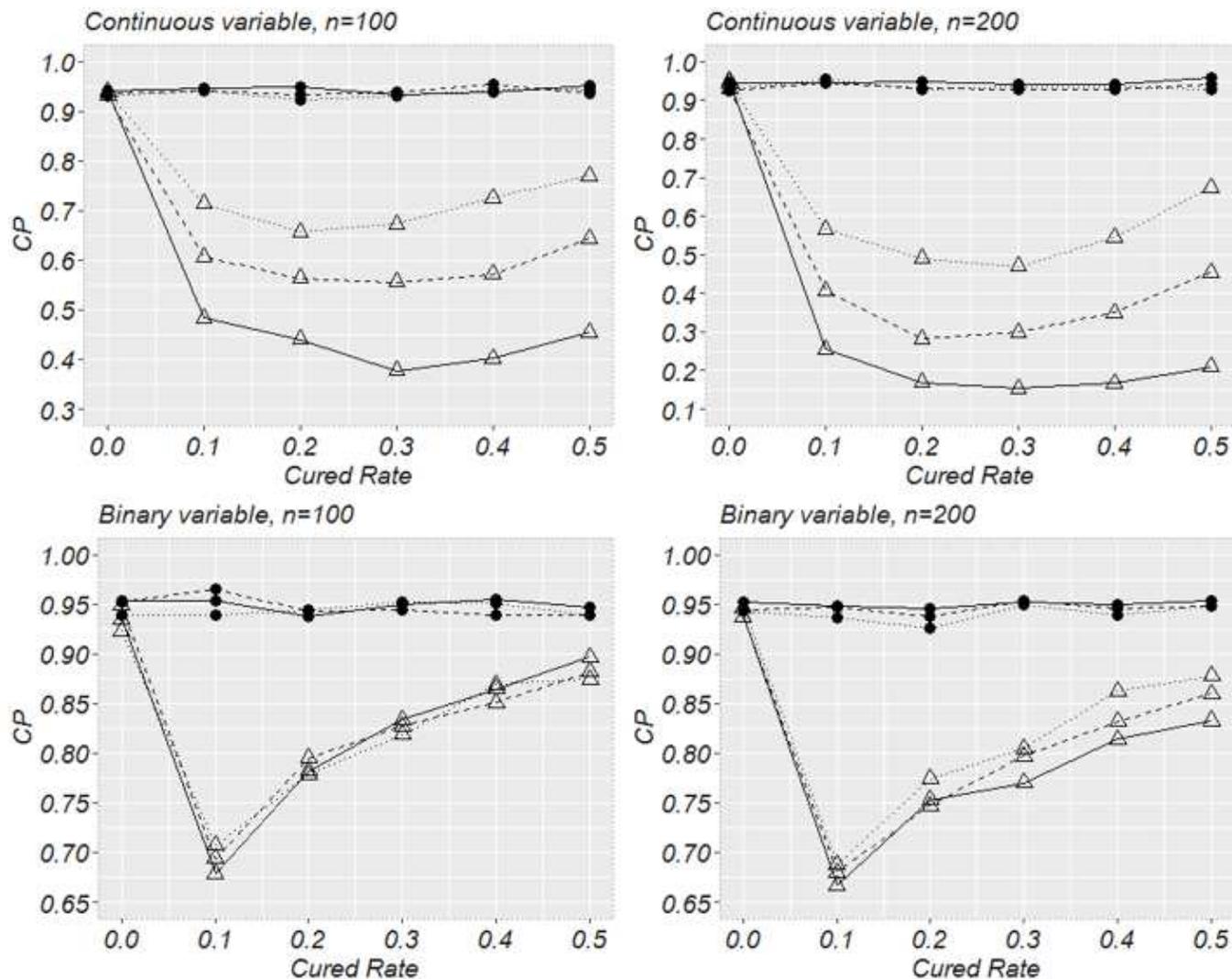


Figure 7

Confidence interval capture rate (CP) for AFT and AFTMC model. Confidence interval capture rate of AFTMC model is its latency part. The triangles represent confidence interval capture rate of AFT model and the circles denote confidence interval capture rate of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

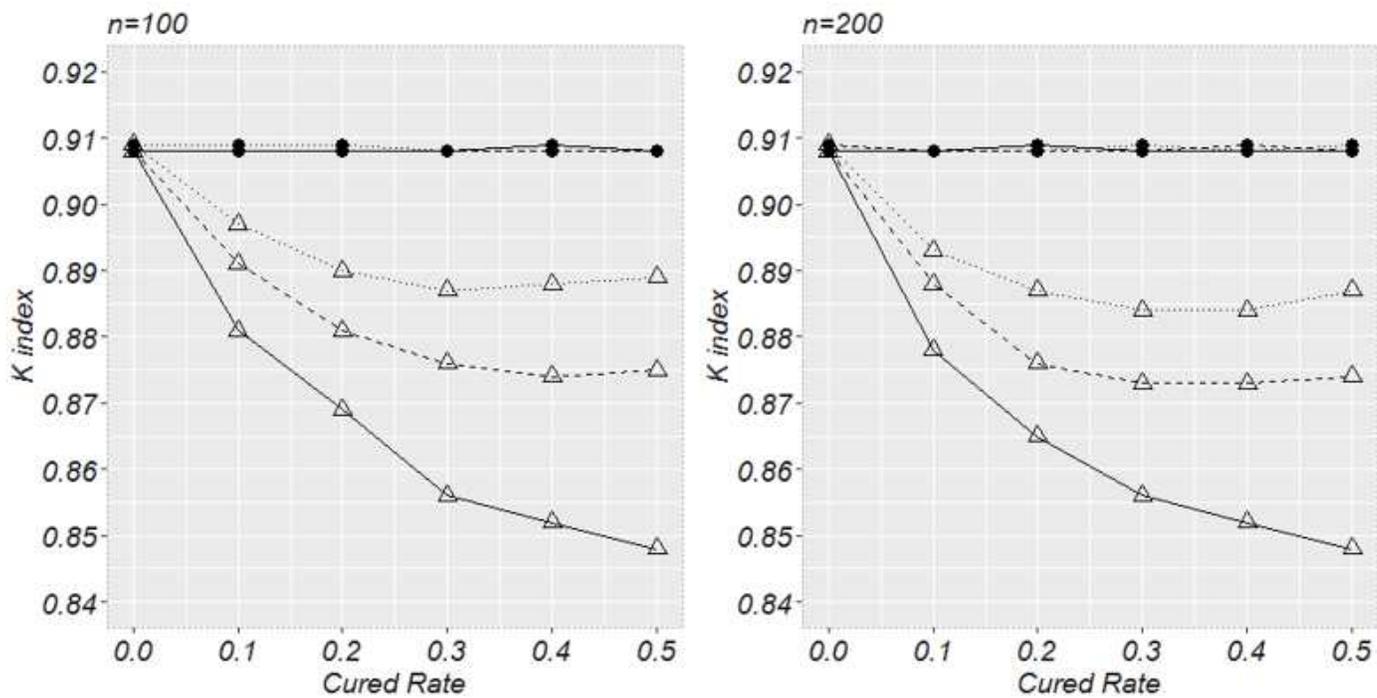


Figure 8

K index for AFT and AFTMC model. The triangles represent the K index of AFT model and the circles denote the K index of AFTMC model. Solid, dashed, and dotted line represent censored rate- cured rate=0.1, 0.2 and 0.3, respectively.

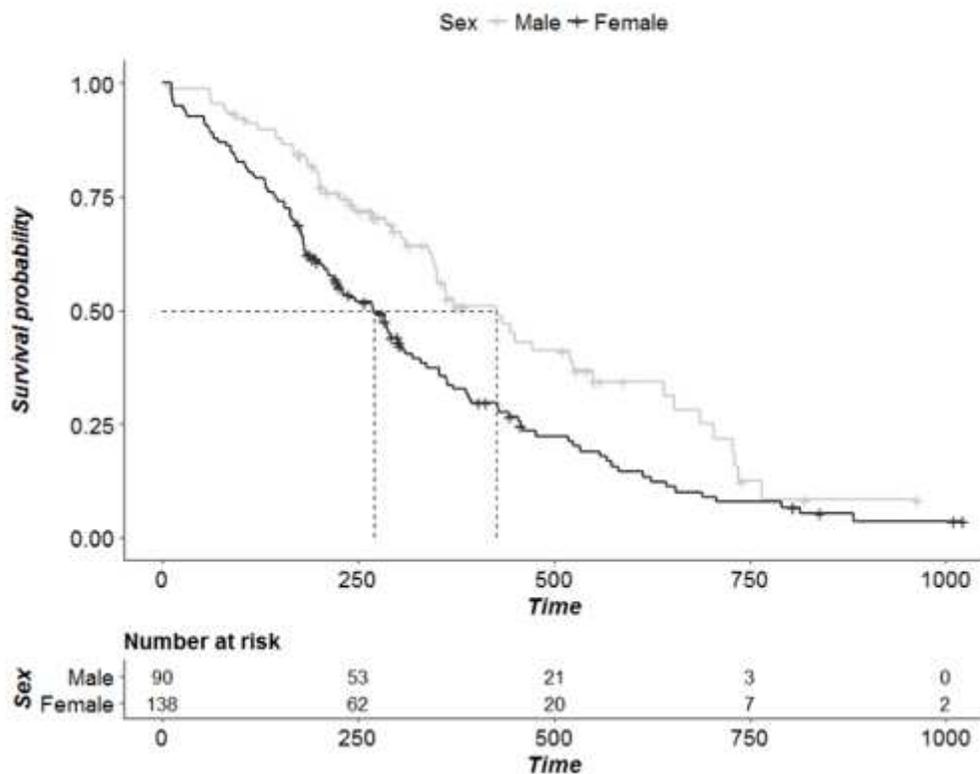


Figure 9

K-M survival curves for lung cancer patients. Time(days)

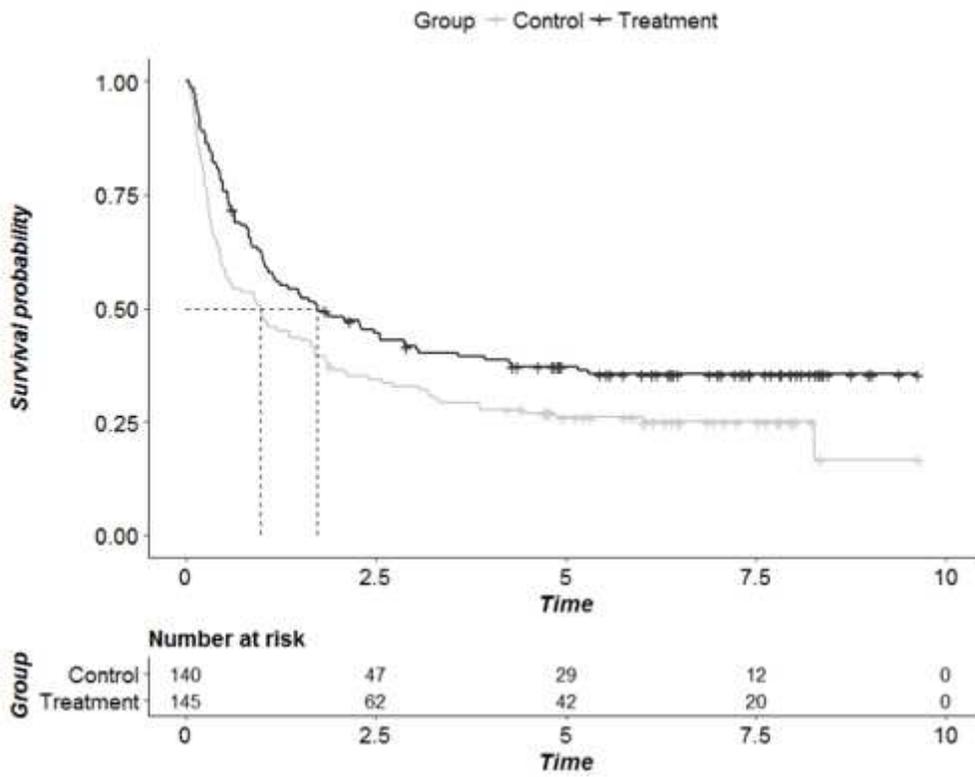


Figure 10

K-M survival curves for E1684 patients. Time(years)