Predictive Biomarker Reproducibility Modeling with Censored Data

by

Qian Kuang

(Under the Direction of Kevin Dobbin)

Abstract

Breast cancer is the most commonly diagnosed cancer within women[1]. A great amount of research has focused on discovering and evaluating predictive biomarkers. In our research, we investigate the interaction between a biomarker and treatment effects(true Θ, which is the decrease in the population event rate under marker-based treatment versus a standard of care)based on the assumption of Cox regression model, and then we conduct a simulation to calculate the estimated Θ under the range of ICC from 0 to 1. We plot the curve of estimated Θ vs. ICC under four different settings. Then we conduct a random effects simulation for the biomarker Ki67, and get the ICC of biomarker Ki67. We conclude that the biomarker is better to detect the treatment effect when the ICC value is greater. We could get the true value of risk rate decrease under marker-based treatment of particular biomarker if we know the estimated value and the ICC of the biomarker in experiments. Our study is informative to evaluate the predictive biomarker detection of treatment effects in cancer.

Index words: Breast Cancer, Predictive Marker, Oncotype DX, TAILORx, Survival Analysis, ICC, Cox Proportional Hazard, Random Effects

Predictive Biomarker Reproducibility Modeling with Censored Data

by

Qian Kuang

B.M., Sun Yat-sen University, 2010

M.S., The University of Texas Pan-American, 2013

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

Master of Science

Athens, Georgia

2016

Predictive Biomarker Reproducibility Modeling with Censored Data

by

Qian Kuang

Approved:

Major Professor: Kevin Dobbin

Committee: Lynne Seymour
Paul Schliekelman

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2016

ACKNOWLEDGMENTS

I will always be grateful to Dr. Dobbin, chair of my thesis committee, for all his mentoring and advice. From research design, and data processing, to manuscript editing, he encouraged me to complete this process through his infinite patience and guidance. My thanks go to my thesis committee Dr. Seymour, and Dr. Schliekelman. Their advice, input, and comments on my thesis helped to ensure the quality of my intellectual work.

I would like to thank Dr. McShane for the experimental data and suggestions. I would also like to thank my colleagues at the UGA library who helped me locate supporting documents for my research. Also, I would like to acknowledge the many volunteers who participated in the focus group research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Breast cancer, which develops from breast tissue, is one of the most dangerous diseases, and is the most commonly diagnosed cancer within women[1]. It affects approximately 12% of women in the world [1]. Since breast cancer is widely distributed worldwide, a lot of treatment options have been developed, such as surgery, radiation oncology, medical oncology, and so on [2].

A great amount of research has been focused on discovering and evaluating markers which can predict a patient's probability of responding to treatment [3]. The treatment selection markers, sometimes called predictive markers, can be used to identify subpopulations of patients who are most likely to respond to a given treatment. The individual patient with predictive biomarkers is able to select the treatment with the best outcome [4]. A predictive marker is a factor which is indicative of sensitivity or resistance to a treatment, and it is important in treating cancer since different cancers have different responses on different treatments. Thus, only some patients will respond to a treatment, while most will only suffer from its side effects [5]. Predictive marker might improve patient outcomes and reduce medical costs as it could allow treatment in those subjects who would benefit from treatment, while avoiding treatment in those who would only likely to suffer its side effects and other costs. Thus, it could be used as indicator of the likely benefit to a specific patient from a specific treatment [6].

One example of predictive markers is the 21-gene Oncotype DX[7]. The 21-gene expression assay provides prognostic information, which is independent of those clinicopathologic

features. It could also predict the benefit from adjuvant chemotherapy in estrogen-receptor-positive disease [7]. The 21-gene recurrence score, which is also called Oncotype DX recurrence score, has a scale from 0 to 100. It is derived from the reference-normalized expression measurements in four steps [8]. The score is higher, the risk of recurrence is greater. The trial assigning individualized options for treatment, which is also called TAILORx, was designed to determine whether the test that analyzes the 21-gene expression is associated with risk of recurrence among women with early-stage breast cancer. Such trial could be applied for treatment selection. The patients with a recurrence score of 0 to 10 were assigned to receive endocrine therapy alone, and the patients with scores from 11 to 25 were randomized to the different treatments: either chemotherapy plus endocrine therapy or endocrine therapy alone, and those with score of 26 or higher received chemotherapy plus endocrine therapy [7].

In most cases, the survival time of breast cancer patients is not complete, and it might be censored at certain times, which forms the survival data; therefore the survival analysis is needed. Survival analysis is a series of statistical processes for which the outcome variable is time until an event occurs. In early study, the term "survival analysis" indicated that the event of interest was death[9]. Now the meaning of survival analysis has become wide, and it could be used for time until occurrence of disease, time until equipment failure, time until earthquake, and so on [9]. Censoring is defined when participants' information is not available on time to event, and it is caused by losing to follow up the participants or absence of the outcome event before the trial ends [10]. There are three reasons which lead to censoring. The first case is that a person does not experience the event before the study ends; a second case is that a person is lost to follow-up during the study period; and the third case is that a person is withdrawn from the study due to death (if the death is not the event of interest) or some other reasons [10].

Since markers that can predict treatment efficacy could make great benefit on improving clinical outcomes and decreasing medical cost, there is much literature on study designs which

are related to treatment selection markers. It is necessary to design a study to evaluate the benefit of a predictive biomarker. However, most research focuses on powering studies which test the statistical interaction between the marker and treatment, and this is not sufficient to evaluate marker performance [11, 12]. An interaction may exist, but the marker not be useful. The positivity criterion, which is a biomarker-based rule for treatment assignment existing before the study, is often uncertainty. In all, the existing trial design method is limited[13].

Based on the reality of survival data, Holly Janes et al. [13] focused on assessing the clinical impact of using the biomarker to assign the treatment, and they used a continuous marker's data to identify a positivity criterion, and evaluated the marker's performance for treatment selection. First, they set two treatment options, binary clinical outcome of interest and the marker value Y. They considered the breast cancer treatment context, proposed the marker-positivity criterion, and listed some statistical measures to evaluate the performance of this treatment rule. One of them is Θ, which is denoted as the difference between the probability of surviving k-years under an optimally-guided-biomarker therapy and under the current standard of care therapy. Then they described two marker evaluation studies. One is called Randomized Control Trial (RCT) design. In this study design, they set four criterions, and described the advantages and disadvantages of each criterion. They used the Oncotype DX marker to illustrate this design by measuring the four criteria and sample size needed to satisfy each of the design criteria. In our research, we applied this type of marker evaluation. For the detailed design method, we will discuss in the next section.

In real laboratory environment, there exist measurement errors of biomarker results, thus it is necessary to conduct reliability studies using intraclass correlation coefficient (ICC) as an index of reliability. There are different approaches to define and classify ICC. If the two-way model is assumed, we have the agreement ICC, which is different from the consistency ICC[14]. The agreement ICC is the ratio which is calculated by the subject variance divided by the sum of the subject variance, the rater variance and the error variance, while the

consistency ICC is the ratio which is calculated by the subject variance divided by the sum of the subject variance and the error variance. The agreement ICC should be used if the variability due to raters is relevant. Another approach is classifying ICC as the following two types: one is called the ICC for single-rater reliability and the other is the ICC for average-rater reliability. The ICC for single-rater reliability is often used to investigate the reliability of individual rating, while the ICC for average-rater reliability is used to evaluate the reliability of average of multiple ratings. In this research, the agreement and single-rater ICC is applied, and the ICC is denoted as: $ICC = \frac{\sigma_t^2}{\sigma_t^2+\sigma_r^2+\sigma_e^2}$, where σ_t^2 is the subject variance, and σ_r^2 is the rater variance, and σ_e^2 is the error variance [15].

In our research, we tried to conduct a simulation to get the association between estimated Θ and ICC. We generated pseudo survival data using Cox proportional hazard. We will discuss this in more detailed in the next section. There are several approaches to analyze time-to-event curves, such as Cox proportional hazards, log-rank, and accelerated failure time models. The Cox proportional hazards model has been widely used [16]. The Cox model is a regression method used for analysis of survival data and identifying differences in survival due to treatment. It could estimate the hazard ratio along with its confidence interval. The Cox regression is also considered as a 'semi-parametric' process because the baseline hazard function does not need to be specified [17].

CHAPTER 2

METHOD

2.1 Setting and Notation

Following notation similar to Janes et al.'s paper[3], let Y_1 be the biomarker value. Let $T = 0$ indicate an individual on a control regimen, and $T = 1$ indicate an individual on a treatment regimen. Let $D = 1$ if an individual dies before a prespecified time t_0 , and $D = 0$ if the individual dies after t_0. This prespecified time is the same for all patients. All data analysis were conducted using the R language version 3.3.0.

2.2 Calculating true Θ

Unlike Janes et al.'s paper[3], we assume a Cox regression model with hazard function

$$h(t) = h_0(t)Exp[\beta_1 Y_1 + \beta_2 T + \beta_3 T Y_1] = Exp[\alpha_0 + \beta_1 Y_1 + \beta_2 T + \beta_3 T Y_1].$$

Here $h_0(t)$ is the baseline hazard, and α_0=$Ln(h_0(t))$. The baseline hazard is exponential. Then the survival function is

$$\begin{aligned} s(t) &= \{Exp[-H_0(t)]\}^{Exp(\beta_1 Y_1+\beta_2 T+\beta_3 T Y_1)} \\ &= \{Exp[-t * Exp(\alpha_0)]\}^{Exp(\beta_1 Y_1+\beta_2 T+\beta_3 T Y_1)} \\ &= Exp[-t * Exp[\alpha_0 + \beta_1 Y_1 + \beta_2 T + \beta_3 T Y_1]], \end{aligned}$$

where $H_0(t)$ is the cumulative hazard function of $h_0(t)$.

When $T = 0$, we get the survival function at t_0 as follows:

$$s(t_0|T = 0) = Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1]].$$

We want to set a t_0 to construct the proportion of death in control regimen so that approximately half of the death times were before t_0 and half were after t_0 in the control regimen, given the parameters α_0 and β_1. Thus, we could avoid having a t_0 with very few events before or after it.

$$P(D = 0|T = 0) = s(t_0|T = 0) = Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1]] = 0.5.$$

Then, we could get

$$t_0 = -Ln(0.5)Exp[-(\alpha_0 + \beta_1 Y_1)].$$

From this, we could get a general formula:

$$t_0 = -Ln(p)Exp[-(\alpha_0 + \beta_1 Y_1)].$$

Here p is the percentile of $D = 0$ under the control regimen. Then we could get

$$P(D = 0) = s(t_0) = Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1 + \beta_2 T + \beta_3 T Y_1]].$$

From this it follows that,

$$\begin{aligned} & P(D = 0|T = 1) - P(D = 0|T = 0) \\ = \;& Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1 + \beta_2 + \beta_3 Y_1]] - Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1]]. \end{aligned}$$

From Janes et al.'s paper[3], let

$$\Delta(Y_1) = P(D = 1|T = 0, Y_1) - P(D = 1|T = 1, Y_1)$$

be the absolute treatment effect given a marker value Y_1. I will make some transformation of this formula as follows:

$$\begin{aligned}
\Delta(Y_1) &= P(D = 1|T = 0, Y_1) - P(D = 1|T = 1, Y_1) \\
&= 1 - P(D = 0|T = 0, Y_1) - (1 - P(D = 0|T = 1, Y_1)) \\
&= P(D = 0|T = 1, Y_1) - P(D = 0|T = 0, Y_1).
\end{aligned}$$

For individuals whose $\Delta(Y_1) < 0$, the optimal way is to recommend against treatment, and for those whose $\Delta(Y_1) > 0$, it's good to recommend treatment.

Then, we could get

$$\begin{aligned}
&P(D = 0|T = 1, \Delta(Y_1) < 0) - P(D = 0|T = 0, \Delta(Y_1) < 0) \\
&= \int_{\Delta(Y_1)<0} \{Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1 + \beta_2 + \beta_3 Y_1]] - Exp[-t_0 Exp[\alpha_0 + \beta_1 Y_1]]\} f(Y_1) dY_1 \\
&= -B_{neg}.
\end{aligned}$$

where B_{neg} is the average benefit of foregoing treatment when $\Delta(Y_1) < 0$ in Janes et al.'s paper[3]. Next, we will rewrite B_{neg} in the formula of death instead of survival.

$$\begin{aligned}
B_{neg} &= P(D = 0|T = 0, \Delta(Y_1) < 0) - P(D = 0|T = 1, \Delta(Y_1) < 0) \\
&= (1 - P(D = 1|T = 0, \Delta(Y_1) < 0)) - (1 - P(D = 1|T = 1, \Delta(Y_1) < 0)) \\
&= P(D = 1|T = 1, \Delta(Y_1) < 0) - P(D = 1|T = 0, \Delta(Y_1) < 0).
\end{aligned}$$

As noted in Janes et al.'s paper[3], P_{neg} is defined as the proportion of individuals that can forego treatment, which could be written as:

$$P_{neg} = P(\Delta(Y_1) < 0).$$

Θ is denoted as the decrease in the population event rate under marker-based treatment. Thus, we could calculate true Θ as follows:

$$\begin{aligned}
\Theta &= [P(D=1|T=1)] - [P(D=1, \Delta(Y_1) \geq 0|T=1)] - [P(D=1, \Delta(Y_1) < 0|T=0)] \\
&= [P(D=1, \Delta(Y_1) < 0|T=1)] - [P(D=1, T=0|\Delta(Y_1) < 0)P(\Delta(Y_1) < 0)] \\
&= [P(D=1, T=1|\Delta(Y_1) < 0)P(\Delta(Y_1) < 0)] - [P(D=1, T=0|\Delta(Y_1) < 0)P(\Delta(Y_1) < 0)] \\
&= [P(D=1, T=1|\Delta(Y_1) < 0) - P(D=1, T=0|\Delta(Y_1) < 0)]P(\Delta(Y_1) < 0) \\
&= B_{neg}P_{neg}.
\end{aligned}$$

Here $B_{neg}P_{neg}$ is the cumulative benefit of foregoing treatment when $\Delta(Y_1) < 0$.

2.3 Simulation

2.3.1 Simulation Overview

In order to look at the bias of the statistical procedure, we did a simulation which generated pseudo survival data. First, we constructed measured biomarker value considering lab effect and error. Then, we used Cox proportional hazard model to generate the survival time. After we got the survival time, we produced two status: Alive and Death. Finally, we applied Holly Janes et. al's approach[3] to calculate the estimated Θ. We also set ICC from 0 to 1 step by 0.05, and plot the association between estimated Θ and ICC.

2.3.2 Construct Measured Biomarker Value

Before we generated pseudo survival data, we denoted

$$Biomarker = Bioleffect + Labeffect + erreffect$$

as the measured biomarker value. Here Bioleffect is the biomarker effect on different subjects(patients), which is set to be normally distributed with mean 0.5 and variance σ_b^2. Labeffect is the effect of each different lab, which is set to be normally distributed with mean 0 and variance σ_l^2. erreffect is the random error, which is set to be normally distributed with mean 0 and variance σ_e^2. All of these effects are independent.

2.3.3 Generate Pseudo Survival Data

As noted in the methods section, the survival time could be generated using the following formula:

$$h(t) = Exp[\alpha_0 + \beta_1 Y_1 + \beta_2 T + \beta_3 T Y_1].$$

Here, Y_1 is the Bioeffect. The survival time is generated from an exponential distribution with rate=h(t).

Next, we denote the status as "Alive" if survival time $> t_0$, and the status as "Death" if survival time $\leq t_0$.

2.3.4 Plot estimated Θ and ICC

Based on the approach in Holly Janes et al. 's paper [3], we could get the estimated Θ. ICC could be computed as the following formula:

$$ICC = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_l^2 + \sigma_e^2}.$$

In the simulation part, our ICC is set from 0 to 1 step by 0.05. For each ICC value, we simulated 2000 times, and got the average estimated Θ.

Finally, we plot a figure which indicates the association between ICC and estimated Θ.

It's worth noting that we use the average estimated Θ of 2000 simulation results, and the 2000 estimated Θ might be truncated at zero. Thus, in some cases, it is necessary to plot the histogram of the 2000 estimated Θ value.

2.3.5 Extension of the Approach

In the previous method, we do not take into consideration that some of the patients are censored. It is possible that the patients' censoring time is less than t_0, but in fact their survival time is longer than t_0.

In the real world, different patients have different censoring time. Thus, we conducted another simulation method by generating a group of survival time s_i and a group of threshold w_i. We denote the status as "censor" if survival time $s_i > w_i$, and the status as "die" if $s_i \leq w_i$.

Since Janes' method does not use the full survival data, but instead it makes survival binary. If we use t_0 to divide the survival time into two status "Alive" and "Death", the patients we recorded "die" is known exactly whether it is survival or not at the time point t_0, but the patients we recorded "censor" might die before or after t_0. If the censor time is longer than t_0, which means the patients is still "Alive" at t_0, but the patients who is censored before t_0 might be "Alive" or "Death" at t_0. In this case, we need to delete the censored cases which had threshold before t_0, and then apply Janes' method. Thus, we could get the estimated Θ in this situation.

2.4 Computing the estimated ICC from Ki67 Data

2.4.1 Background of Ki67 Data

The Ki67 data was collected using the same method from Polley et al.'s paper[18],

There were 100 breast cancer cases, and they were arranged into 1-mm core tissue microarrays(TMAs), and each 50 of them represented on each of two TMA blocks. There were eight laboratories from North America and Europe participating in this experiment. Each tissue would be stained using the method introduced in Polley et al's paper[18], and the percentage of tumor cells positively stained was recorded as the final Ki67 score. We have two groups of data. One is called "experiment A", and another is called "experiment B".

2.4.2 Statistical Analysis of Ki67 Data

There was very little missing data in the experiments (0.75% in experiment A and 2.38% in experiment B), and the missing data was dealt with filling them with the average of each patient's Ki67 score. We would like to apply random effects model in this research. Since the random effects model requires the data being normally distributed with constant variance, we took a BoxCox-transformation with Ki67 data to make it approximately normally distributed with constant variance.

After dealing with the raw data, the random effects model was fitted using the lme4 package [19].

The random effects model is as follows:

$$Y_{ij} = u + B_i + L_j + \varepsilon_{ij}.$$

Where Y_{ij} is the Ki67 score after BoxCox transformation of the jth lab at the ith subject. u is the avarage score of the entire population. B_i is the biological subject-specific random effect. L_j is the lab-specific random effect.

2.4.3 Computing the estimated ICC from $\hat{\sigma_b^2}$, $\hat{\sigma_l^2}$, and $\hat{\sigma_e^2}$

We could get the $\hat{\sigma_b^2}$, $\hat{\sigma_l^2}$, and $\hat{\sigma_e^2}$ value from the R output of random effects model, which is shown in the Appendix B.1. We could also calculate the estimated ICC of Ki67 data using this formula:

$$\widehat{ICC} = \frac{\hat{\sigma_b^2}}{\hat{\sigma_b^2} + \hat{\sigma_l^2} + \hat{\sigma_e^2}}.$$

CHAPTER 3

RESULTS

3.1 Simulation Result

3.1.1 Parameter Settings

Before we conducted simulation, we have these following four groups of parameter setting shown in **Table** 3.1.

λ	β_1	β_2	β_3	$-\beta_2/\beta_3$
0.6086	-3.145	-3.145	6.290	0.5
0.2454	-1.169	-1.169	2.337	0.5
0.0262	3.145	3.145	-6.290	0.5
0.1386	0.000	1.479	-3.145	0.5

Table 3.1: Four groups of parameter setting

3.1.2 Result Plots

If we set λ =0.6086, β_1=-3.145, β_2=-3.145, β_3=6.290, we could get a true Θ=0.2137063. The biomarker is normally distributed with mean 0.5 and variance 1/12.

In this parameter setting group, we have the risk curve which is shown in **Figure** 3.1a. It has a treatment risk line and a non-treatment risk line. The intersection between the treatment line and non-treatment line is pointing at around 44% population below marker value around 0.47.

The plot of estimated Θ vs. ICC is shown in **Figure** 3.1b. From the figure, we could find that slope=Θ*1=Θ, and then estimated Θ=ICC*Θ =0.2137063*ICC.

To compare easily, we also plot the estimated Θ vs. ICC when there is a censored problem (We generated censored data and eliminated the ones with censored time before the t_0 cutoff) which is discussed in the extension part of method. This plot is shown in **Figure** 3.1c. The simulation that goes with 3.1c, 3.2c, 3.3c, and 3.4c contains about 30% of censoring data before we deleted the ones with censored time ahead t_0, but it contains only about 15% after we deleted them. The percentage that were deleted during analysis is roughly 18%.

If we set λ =0.2454, β_1=-1.169, β_2=-1.169, β_3=2.337, we could get a true Θ=0.09105615. The biomarker is normally distributed with mean 0.5 and variance 1/12.

In this parameter setting group, we have the risk curve which is shown in **Figure** 3.2a. The intersection between the treatment line and non-treatment line is pointing at around 60% of population below marker value around 0.58.

The plot of estimated Θ vs. ICC is shown in **Figure** 3.2b. From the figure, we could find estimated Θ=ICC*Θ =0.09105615*ICC.

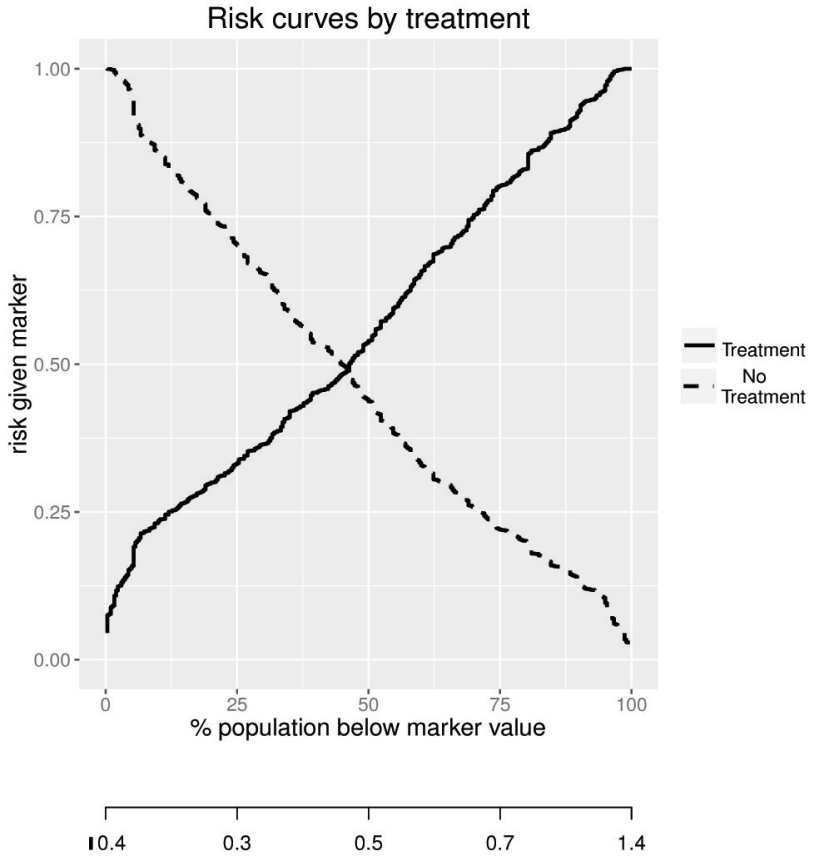
To compare easily, we also plot the estimated Θ vs. ICC when we generated censored data and eliminated the ones with censored time before the t_0 cutoff. This plot is shown in **Figure** 3.2c.

If we set λ =0.0262, β_1=3.145, β_2=3.145, β_3=-6.290, we could get a true Θ= 0.213746. The biomarker is normally distributed with mean 0.5 and variance 1/12.

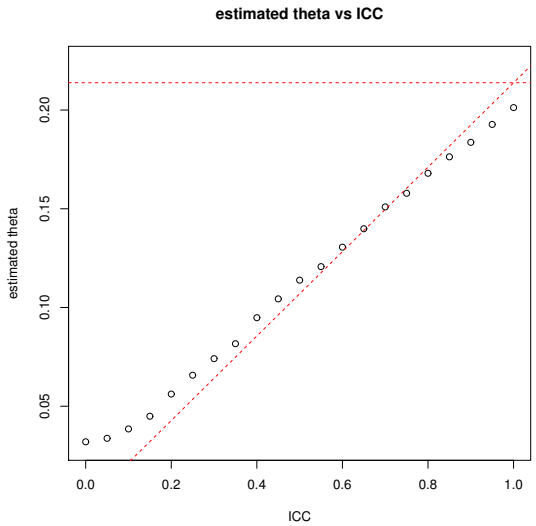
In this parameter setting group, we have the risk curve which is shown in **Figure** 3.3a. The intersection between the treatment line and non-treatment line is pointing at around 49% of population below biomarker value around 0.49.

The plot of estimated Θ vs. ICC is shown in **Figure** 3.3b. From the figure, we could find estimated Θ=ICC*Θ =0.213746*ICC.

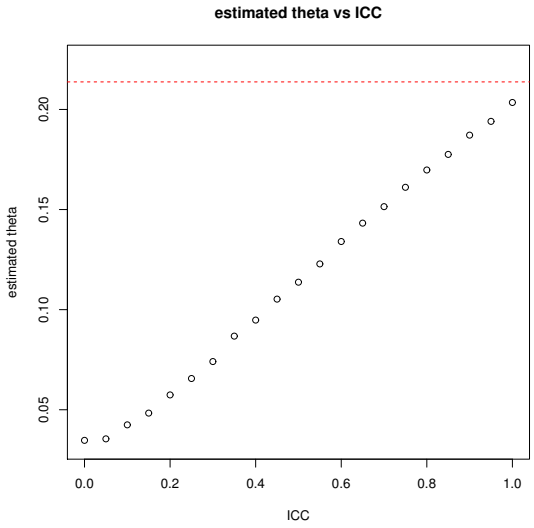
To compare easily, we also plot the estimated Θ vs. ICC when we generated censored data and eliminated the ones with censored time before the t_0 cutoff. This plot is shown in **Figure** 3.3c.



(a) risk curve

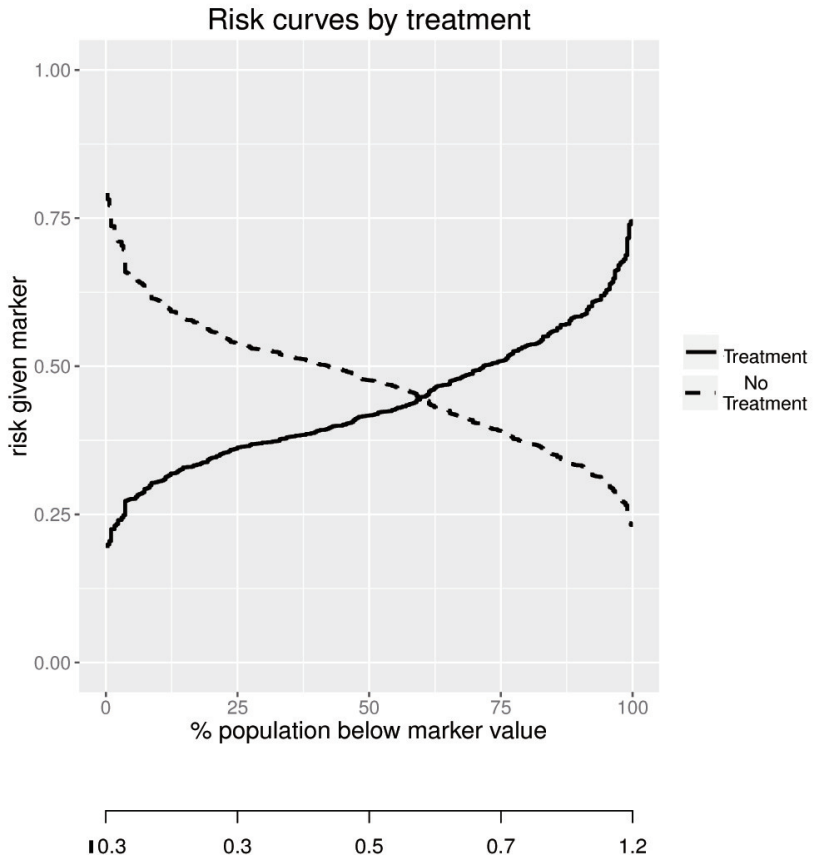


(b) estimated Θ vs. ICC

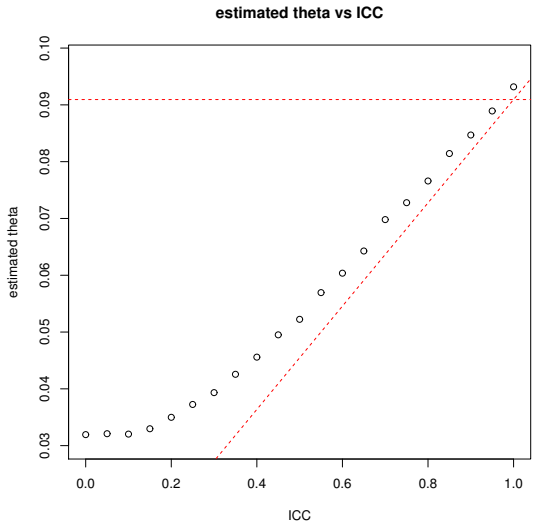


(c) estimated Θ vs. ICC with censored problem

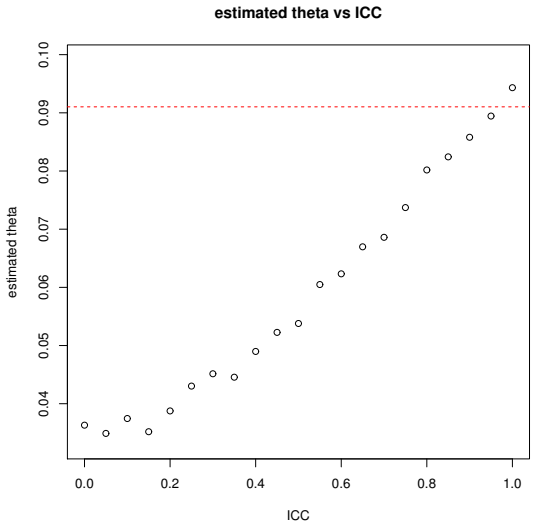
Figure 3.1: group 1 of parameter settings. In 3.1(a), the biomarker values are shown on the lowest axis.



(a) risk curve

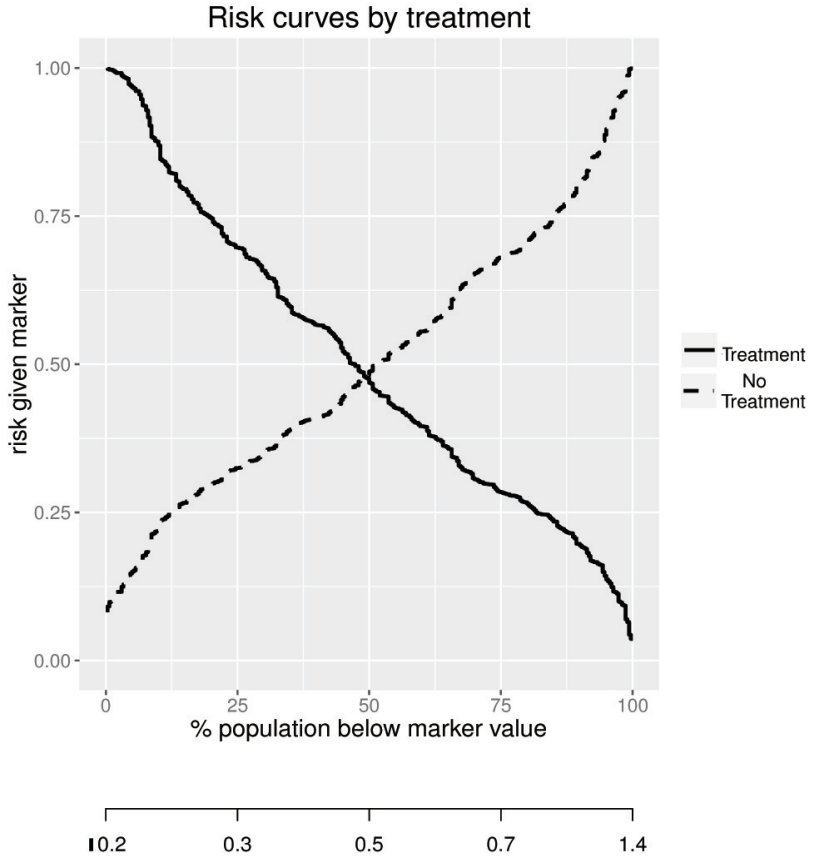


(b) estimated Θ vs. ICC

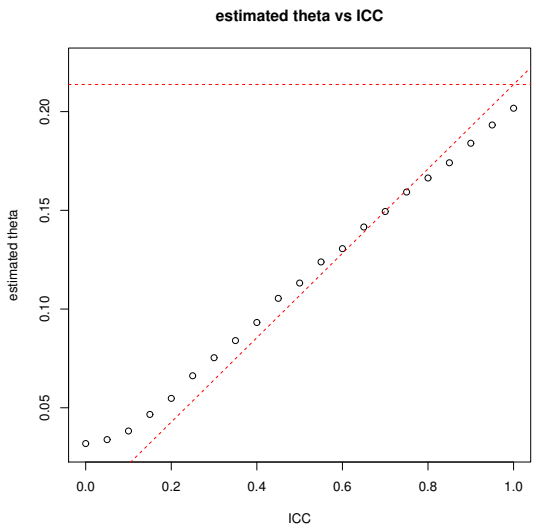


(c) estimated Θ vs. ICC with censored problem

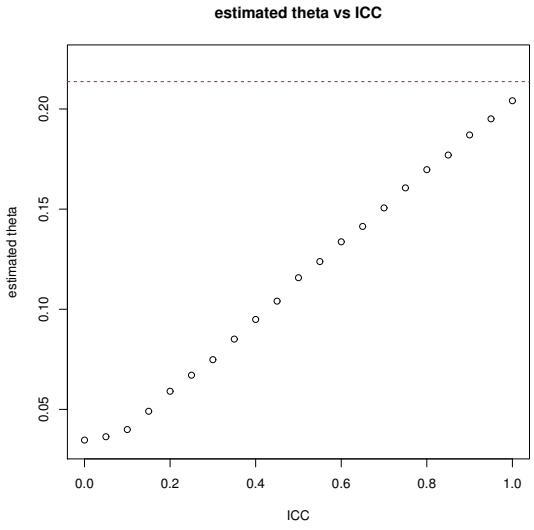
Figure 3.2: group 2 of parameter settings. In 3.2(a), the biomarker values are shown on the lowest axis.



(a) risk curve



(b) estimated Θ vs. ICC



(c) estimated Θ vs. ICC with censored problem

Figure 3.3: group 3 of parameter settings. In 3.3(a), the biomarker values are shown on the lowest axis.

If we set λ =0.1386, β_1=0.000, β_2=1.479, β_3=-3.145, we could get a true Θ= 0.1036687. The biomarker is normally distributed with mean 0.5 and variance 1/12.

In this parameter setting group, we have the risk curve which is shown in **Figure** 3.4a. The intersection between the treatment line and non-treatment line is pointing at around 42.5% of population below biomarker value around 0.45.

The plot of estimated Θ vs. ICC is shown in **Figure** 3.4b. From the figure, we could find estimated Θ=ICC*Θ =0.1036687*ICC.

To compare easily, we also plot the estimated Θ vs. ICC when we generated censored data and eliminated the ones with censored time before the t_0 cutoff. This plot is shown in **Figure** 3.4c.

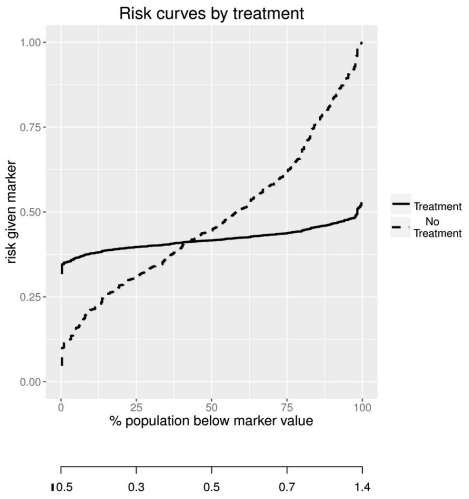
In this case, we set ICC=0 and ICC=0.5, and plot the histogram of the 2000 estimated Θ. The histogram of estimated Θ when ICC=0 is shown in **Figure** 3.5a, and the histogram of estimated Θ when ICC=0.5 is shown in **Figure** 3.5b.

I also changed sample length from 300 to 1000, and plot the estimated Θ vs. ICC, and the figure is shown in **Figure** 3.4d. If I increase the proportion of censoring data from around 30% to around 70% before we deleted the ones with censored time ahead t_0, we have the plot of estimated Θ vs. ICC in **Figure** 3.4e.

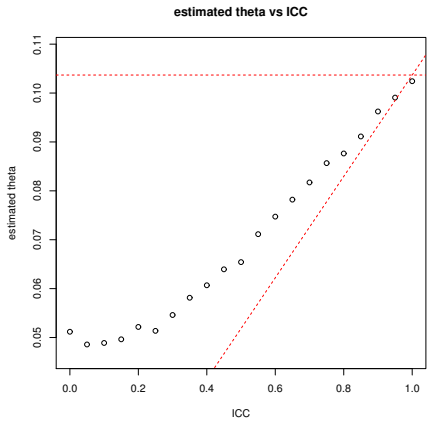
3.2 Real Data Result

First, we get the scatter plot of the original data. The scatter plot of the original data of experiment A is shown in **Figure** 3.6, and the scatter plot of B is shown in **Figure** 3.7. From these two plots, we could find most of these correlation is positive with non-constant variance. The variance is increasing when the Ki67 score is increasing. Thus, it might be necessary to conduct a BoxCox transformation.

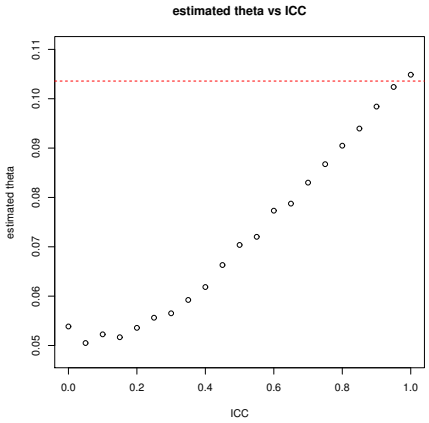
After we conducted a BoxCox test, we get that the λ of BoxCox transformation of experiment A is 0.2614825, and the λ of BoxCox transformation of experiment B is 0.1846809. Both of the BoxCox confidence interval of λ do not contain 1. Thus, it is necessary to conduct



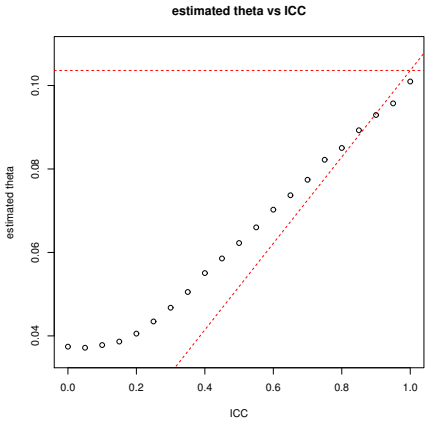
(a) risk curve



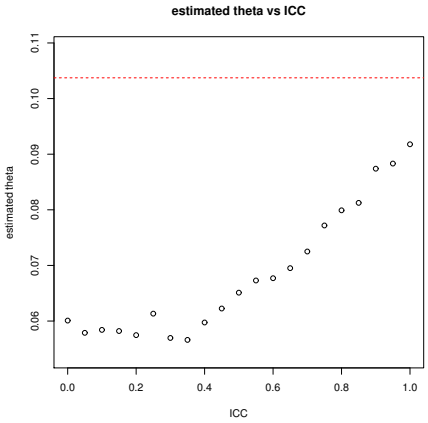
(b) estimated Θ vs. ICC when sample length=300



(c) estimated Θ vs. ICC with censored problem

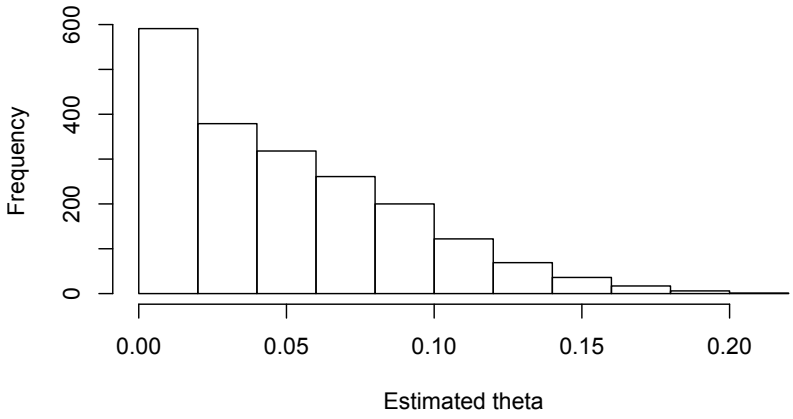


(d) estimated Θ vs. ICC when sample length=1000

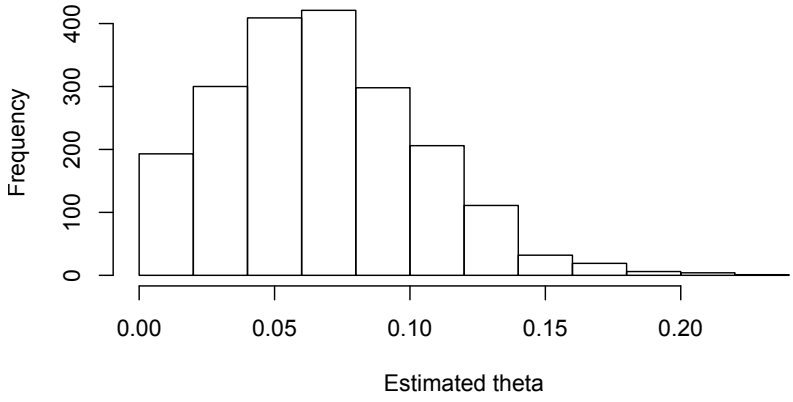


(e) estimated Θ vs. ICC with a lot of censoring

Figure 3.4: group 4 of parameter settings. In 3.4(a), the biomarker values are shown on the lowest axis.



(a) Histogram of estimated Θ when ICC=0



(b) Histogram of estimated Θ when ICC=0.5

Figure 3.5: Histogram of estimated Θ

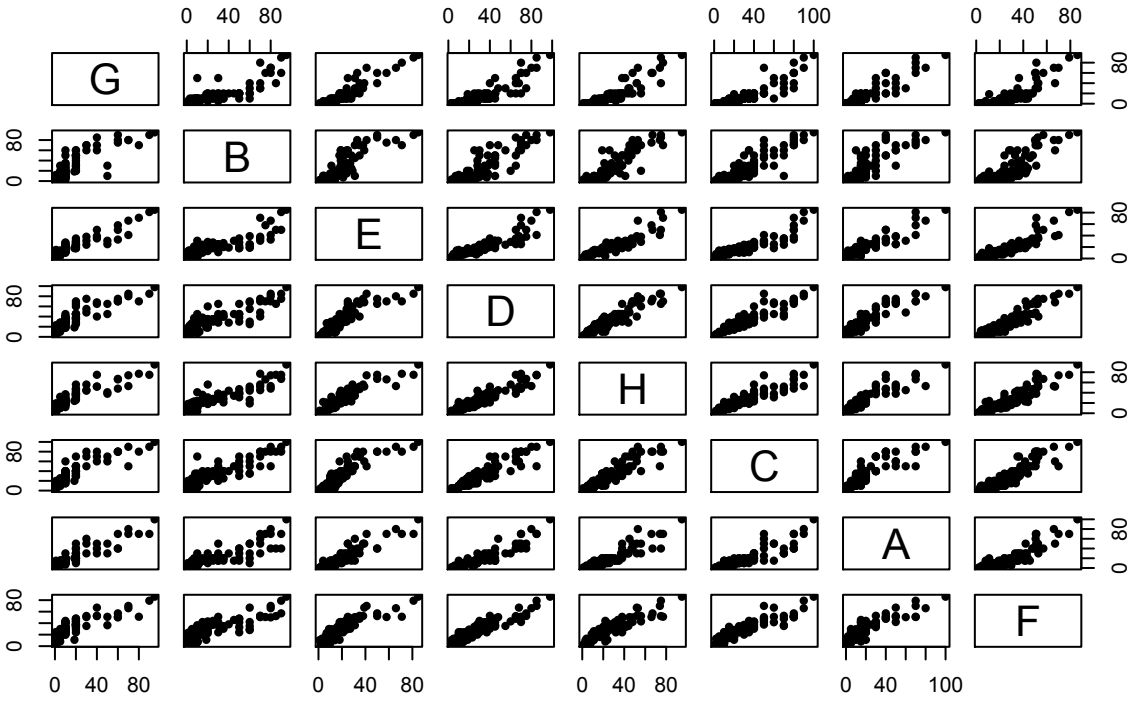


Figure 3.6: scatterplot of original data in experiment A

a BoxCox transformation. After the data was transformed, we drew the scatter plot of each experiment again. The scatter plot of A is shown in **Figure** 3.8, and the scatter plot of B is shown in **Figure** 3.9. The scatterplot shows that the Ki67 score between every two labs are positive correlated with constant variance, though there exists some outliers. Then we could use the data after transformation to fit the random effects model.

After we fitted a random effect model, we get the following results of experiment A: $\hat{\sigma_b^2}$=3.4951, $\hat{\sigma_l^2}$=0.5830, $\hat{\sigma_e^2}$=0.5774.

Thus we could calculate the estimated ICC of experiment A:

$$\widehat{ICC} = \frac{\hat{\sigma_b^2}}{\hat{\sigma_b^2} + \hat{\sigma_l^2} + \hat{\sigma_e^2}} = \frac{3.4951}{3.4951 + 0.5830 + 0.5774} = 0.7507464$$

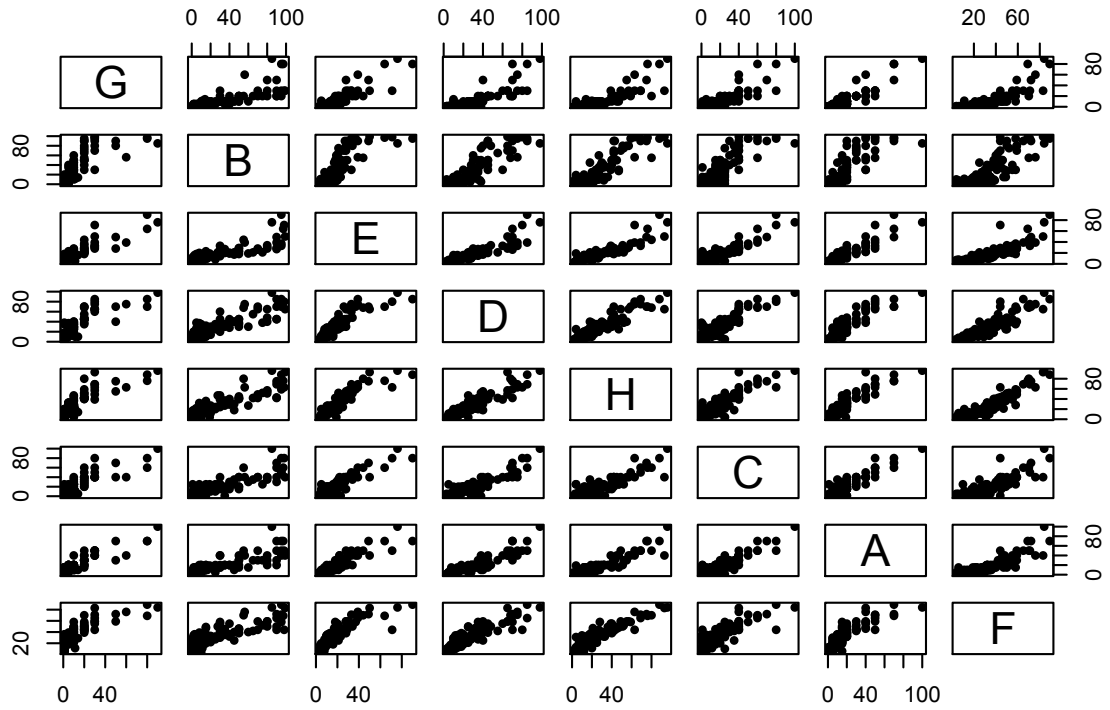


Figure 3.7: scatterplot of original data in experiment B

We get the following results of experiment B: $\hat{\sigma_b^2}$=2.2415, $\hat{\sigma_l^2}$=0.5734, $\hat{\sigma_e^2}$=0.8439.

Thus, we could calculate the estimated ICC of experiment B:

$$\widehat{ICC} = \frac{\hat{\sigma_b^2}}{\hat{\sigma_b^2} + \hat{\sigma_l^2} + \hat{\sigma_e^2}} = \frac{2.2415}{2.2415 + 0.5734 + 0.8439} = 0.6126326$$

In contrast, Janes et al. [3] used a base 2 log transformation which results in an estimated ICC of 0.71 in experiment A and estimated ICC of 0.59 in experiment B.

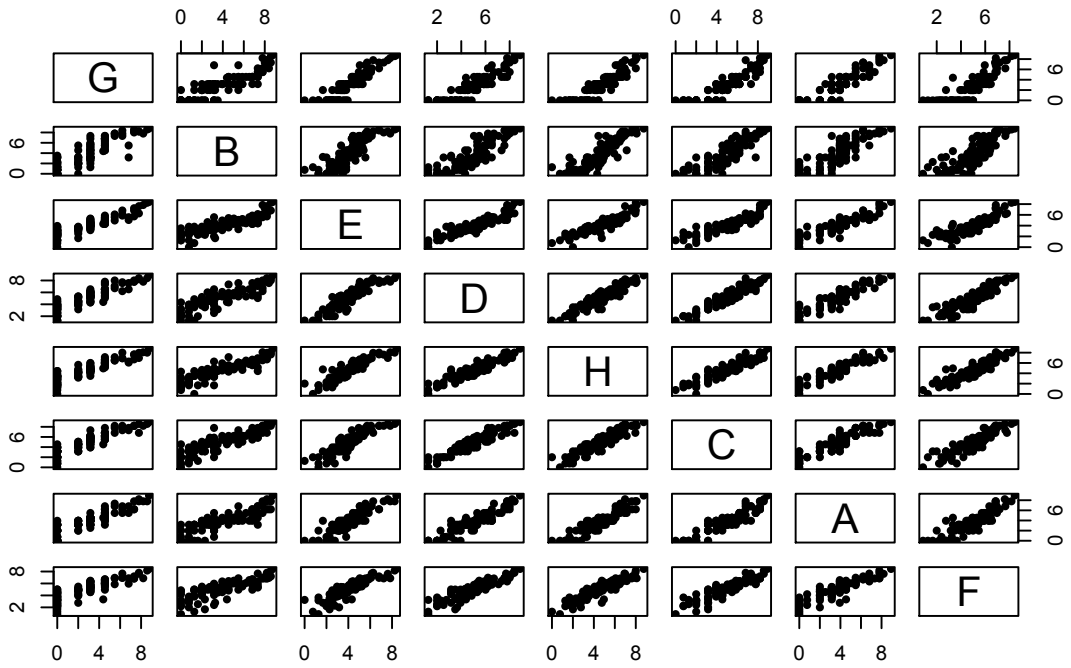


Figure 3.8: scatterplot of experiment A

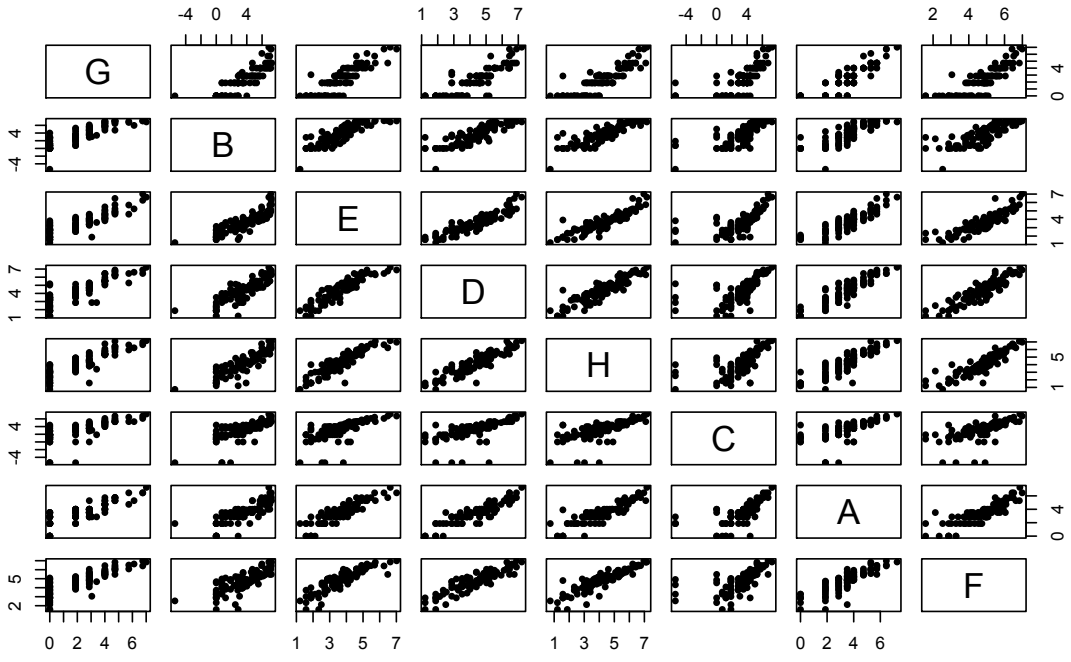


Figure 3.9: scatterplot of experiment B

CHAPTER 4

DISCUSSION

When λ =0.6086, β_1=-3.145, β_2=-3.145, β_3=6.290, the risk curve indicates that this biomarker is very strong since there is a dramatically big difference between the treatment risk and non-treatment risk. This also indicates that the slope of the plot of estimated Θ vs. ICC would be larger. From the risk curve, we could conclude that the proportion of subjects with positive treatment effects who need treatment is around 44% below the marker value about 0.47. The fact that the proportion is estimated to be 44% rather than 50% is just because of the limited dataset(n=300) and there is always some noise in the estimates. This situation also occurs in the other three groups of parameter settings. The slope of the plot of estimated Θ vs. ICC is close to the true Θ value as the dot is almost on the straight line with slope of 0.2137063. From the plot of ICC vs. estimated Θ under the situation with censored problem which we have discussed in the method extension part, we could find the bias is small and the plot looks very similar to the plot without censored problem. In this situation, we have generated the censored data and then eliminated the ones with censored time before t_0 cutoff.

When λ =0.2454, β_1=-1.169, β_2=-1.169, β_3=2.337, the risk curve indicates that this biomarker is not so strong as the biomarker of the first group of parameter settings. But there is still a big difference between the treatment risk and non-treatment risk. This also indicates that the slope of the plot of estimated Θ vs. ICC would be smaller than the first group. From the risk curve, we could conclude that the proportion of subjects with positive treatment effects who need treatment is around 60% below the marker value about 0.58 .

The slope of the plot of estimated Θ vs. ICC is also close to the true Θ value as the dot is almost on the straight line with slope of 0.09105615. After we generated the censored data and then eliminated the ones with censored time before t_0 cutoff, we have the plot of ICC vs. estimated Θ under this situation, and we could find the bias is small and the plot looks very similar to the plot without censored data except that the curve is not so smooth.

When λ =0.0262, β_1=3.145, β_2=3.145, β_3=-6.290, the risk curve indicates that this biomarker is as strong as the biomarker of the first group of parameter settings since there is also a dramatically big difference between the treatment risk and non-treatment risk. This also indicates that the slope of the plot of estimated Θ vs. ICC would be large. From the risk curve, we could conclude that the proportion of subjects with negative treatment effects who should avoid treatment is around 49% below the marker value about 0.49. The slope of the plot of estimated Θ vs. ICC is also close to the true Θ value as the dot is almost on the straight line with slope of 0.213746. After we generated the censored data and then eliminated the ones with censored time before t_0 cutoff, we get another plot of ICC vs. estimated Θ, and we could find that the bias is small and the plot looks very similar to the plot without censored data.

When λ =0.1386, β_1=0.000, β_2=1.479, β_3=-3.145, the risk curve indicates that this biomarker is weak since this plot is not symmetric, and the treatment line is increasing, while the non-treatment line is almost flat. There is not a big difference between the treatment risk and non-treatment risk. This indicates the slope of the plot of estimated Θ vs. ICC would be small. From the risk curve, we could conclude the proportion of subjects with negative treatment effects who should avoid treatment is around 42.5% below the biomarker value about 0.45.

The slope of the plot of estimated Θ vs. ICC is not so close to the true Θ value as the dot is a little away from the straight line with slope of 0.1036687. Thus, we plotted the histogram of the 2000 estimated Θ when ICC=0 and ICC=0.5, and from the figure, we could find that the histogram is truncated at 0. This indicates the curve of estimated Θ vs.

ICC should be close to the straight line with slope of true Θ, and when ICC=0, estimated Θ should also be equal to zero. The plot with estimated Θ not equal to zero is caused by that we have used the average of estimated Θ to plot it, which is not zero any more.

It also indicates that there is a bias in the Janes et al.'s method for values of Θ near 0. This is a new and somewhat interesting finding (not a big deal because usually if Θ is near 0, then the marker is not of interest, but it suggests that the method probably need improvement).

Similar to the other three groups of parameter settings, we get the plot of ICC vs. estimated Θ after generating the censored data and deleting the ones with censored time before t_0. We could find the bias is small and the plot looks very similar to the plot without censored problem. However, if there exists a lot of censoring, there is a little bias of 0.01 between true Θ and estimated Θ.

In the figure with length equal to 1000, we could find that the curve of estimated Θ vs. ICC is closer to the straight line compared to the previous plot, which indicates that the estimated Θ value might be more accurate if we increase the sample length.

From the result of real data Biomarker Ki67, we could find that estimated ICC of experiment A is 0.7507464, and the estimated ICC of experiment B is 0.6126326. Since the estimated ICC could not be one in real world, thus the estimated marker-based treatment effects might be worse than the real treatment effects. From previous discussion, we could suppose the true Θ of Biomarker Ki67 to be around 0.3 to 0.6. For example, if we have the true Θ of Biomarker Ki67 which is 0.6, and it means the risk rate could decrease 60% under treatment. We could draw a straight line of estimated Θ=0.6*$\widehat{ICC}$. Since we could not get the true Θ in real world as the biomarker is not perfect, we could only get the estimated Θ. In experiment A, we could have that the estimated Θ is 0.6*0.7507464 $\approx$ 0.45, which means the risk rate decrease 45% under treatment. This indicates that the marker-based treatment effects decrease from 60% to 45% using the Ki67 biomarker in experiment A. Similarly to experiment B, we have that the estimated Θ is 0.6*0.6126326 $\approx$ 0.37, which means the risk

rate decrease 37% under treatment. This also indicates that the treatment effects decrease from 60% to 37% using the Ki67 biomarker in experiment B. Besides, we could also draw a straight line once we get the estimated Θ, and this line will pass through (0,0). We could also get the estimated ICC value in experiments, thus we could have an idea what the true Θ is, which means that we could get a much more precise estimation of the true treatment effects if we have the coarse estimated treatment effects from experiments.

CHAPTER 5

CONCLUSION

In our research, we have looked at risk curve under four different groups of parameter settings. It shows that some biomarkers are strong and some biomarkers are weak.

We also looked at the association between ICC and estimated Θ. We could conclude that the estimated Θ decreases in a nearly linear fashion (that is probably faster than one might have guessed) when ICC decreases. It also indicates that the Biomarker with greater reliability has a greater risk rate decrease under marker-based treatment.

From the discussion, we could also conclude that the straight line with slope of true Θ is close to the plot of estimated Θ vs. ICC, which also indicates that we could have a better estimation of the true value of risk rate decrease under marker-based treatment for particular biomarker if we have the estimated value of risk rate decrease under marker-based treatment and the reliability of experiments.

We have found that if there is censoring in the data, the bias of the estimator in Jane's method is small unless the censoring proportion gets to be large (e.g. well over 50%). When we used Janes et al.'s method to calculate the estimated Θ, we might have a problem of estimated Θ with truncated zero normal distribution, which produces a positive bias for value of Θ near 0. A better way to get a more exact estimated Θ under simulation might need to be further studied.

REFERENCES

[1] Andrew McGuire, James AL Brown, Carmel Malone, Ray McLaughlin, and Michael J Kerin. Effects of age on the detection and management of breast cancer. *Cancers*, 7(2):908–929, 2015.

[2] Breast Cancer - Treatment Options. http://www.cancer.net/cancer-types/breast-cancer/treatmentoptions, 2016. Online; accessed 14-Aug-2016.

[3] Holly Janes, Marshall D Brown, Ying Huang, and Margaret S Pepe. An approach to evaluating and comparing biomarkers for patient treatment selection. *The international journal of biostatistics*, 10(1):99–121, 2014.

[4] Nils Brünner. What is the difference between "predictive and prognostic biomarkers"? can you give some examples. *Connection*, 13:18, 2009.

[5] Michael J Duffy. Predictive markers in breast and other cancers: a review. *Clinical Chemistry*, 51(3):494–503, 2005.

[6] Lisa M McShane, Douglas G Altman, Willi Sauerbrei, Sheila E Taube, Massimo Gion, and Gary M Clark. Reporting recommendations for tumour marker prognostic studies (remark). *British journal of cancer*, 93(4):387–391, 2005.

[7] Joseph A Sparano, Robert J Gray, Della F Makower, Kathleen I Pritchard, Kathy S Albain, Daniel F Hayes, Charles E Geyer Jr, Elizabeth C Dees, Edith A Perez, John A Olson Jr, et al. Prospective validation of a 21-gene expression assay in breast cancer. *New England Journal of Medicine*, 373(21):2005–2014, 2015.

[8] Soonmyung Paik, Steven Shak, Gong Tang, Chungyeul Kim, Joffre Baker, Maureen Cronin, Frederick L Baehner, Michael G Walker, Drew Watson, Taesung Park, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *New England Journal of Medicine*, 351(27):2817–2826, 2004.

[9] Tyler Smith and Besa Smith. Survival analysis and the application of cox's proportional hazards modeling using sas. In *Proceedings of the twenty-sixth annual SAS user?s group international conference. SAS Institute Inc, Cary, NC*, pages 244–246, 2001.

[10] Shankar Prinja, Nidhi Gupta, and Ramesh Verma. Censoring in clinical trials: Review of survival analysis techniques. *Indian journal of community medicine: official publication of Indian Association of Preventive & Social Medicine*, 35(2):217, 2010.

[11] Holly Janes, Margaret S Pepe, Patrick M Bossuyt, and William E Barlow. Measuring the performance of markers for guiding treatment decisions. *Annals of internal medicine*, 154(4):253–259, 2011.

[12] Ying Huang, Peter B Gilbert, and Holly Janes. Assessing treatment-selection markers using a potential outcomes framework. *Biometrics*, 68(3):687–696, 2012.

[13] Holly Janes, Marshall D Brown, and Margaret S Pepe. Designing a study to evaluate the benefit of a biomarker for selecting patient treatment. *Statistics in medicine*, 34(27):3503–3515, 2015.

[14] Kenneth O McGraw and Seok P Wong. Forming inferences about some intraclass correlation coefficients. *Psychological methods*, 1(1):30, 1996.

[15] Yuki Saito, Takashi Sozu, Chikuma Hamada, and Isao Yoshimura. Effective number of subjects and number of raters for inter-rater reliability studies. *Statistics in medicine*, 25(9):1547–1560, 2006.

[16] David Roxbee Cox and David Oakes. *Analysis of survival data*, volume 21. CRC Press, 1984.

[17] Ritesh Singh and Keshab Mukhopadhyay. Survival analysis in clinical trials: Basics and must know areas. *Perspectives in clinical research*, 2(4):145, 2011.

[18] Mei-Yin C Polley, Samuel CY Leung, Lisa M McShane, Dongxia Gao, Judith C Hugh, Mauro G Mastropasqua, Giuseppe Viale, Lila A Zabaglo, Frédérique Penault-Llorca, John MS Bartlett, et al. An international ki67 reproducibility study. *Journal of the National Cancer Institute*, page djt306, 2013.

[19] CRAN - Package lme4. https://cran.r-project.org/web/packages/lme4/index.html, 2016. Online; accessed 04-Sep-2016.

APPENDIX A

CODE

A.1 Simulation Code

A.1.1 Without Censoring Data

```
rm(list=ls())

#######calculate true Theta###########

normalmean = 0.5;

normalvar = 1/12;

mcreps = 2000;

sample_length <- 300

#sample_length <- 1000

T<-round(runif(sample_length,min=0,max=1))

library(survival)

ICClist = (0:20)/20;

#ICClist=0;

#ICClist=0.5;

sigmasqblist = ICClist/12;

sigmasqelist = (1/24)*(1-ICClist);

sigmasqllist = sigmasqelist;

Bioleffectvec<-rnorm(1e7,mean=normalmean,sd=sqrt(normalvar))
```

```
alpha0 = log(0.6086)

beta1 = -3.145

beta2 = -3.145

beta3 =6.290

plot_index = 1


#alpha0 = log(0.2454)

#beta1 = -1.169

#beta2 = -1.169

#beta3 =2.337

#plot_index = 2


#alpha0 = log(0.0262)

#beta1 = 3.145

#beta2 = 3.145

#beta3 =-6.290

#plot_index = 4


#alpha0 = log(0.1386)

#beta1 = 0

#beta2 = 1.479

#beta3 =-3.145

#plot_index = 5


t0percentile = 0.5;

# This below code sets t0 to be about the approx median of the survival times

# under control, to change, change t0percentile
```

```
t04param = -log(t0percentile)*exp(-(alpha0+beta1*normalmean));
t04param;

# delta_Y = exp(0-t04param*exp(alpha0+beta1*Bioleffectvec+beta2+beta3*Bioleffectvec))
#                 -exp(-t04param*exp(alpha0+beta1*Bioleffectvec))
delta_Y = 1 - exp(-t04param*exp(alpha0))^exp(beta1*Bioleffectvec) -1
+ exp(-t04param*exp(alpha0))^exp(beta1*Bioleffectvec+beta2+beta3*Bioleffectvec);

stat_delta_Y <-NULL
stat_delta_Y[which(delta_Y<0)]=1
stat_delta_Y[which(delta_Y>=0)]=0
P_neg = sum(stat_delta_Y)/length(delta_Y)

B_neg = 0-sum(delta_Y*stat_delta_Y)/length(which(stat_delta_Y==1))
B_neg
theta <-B_neg*P_neg
theta
########calculate estimated Theta and plot ICC-Theta##########
ICC<-NULL
#theta_hat_e<-NULL
#theta_hat_mod<-NULL
library(doParallel)
mycluster = makeCluster(3)
registerDoParallel(mycluster)
getDoParWorkers()
library(foreach)
result1 <- foreach(j = 1:length(ICClist),.combine=rbind) %dopar% {
```

```
library(TreatmentSelection)
ICC = ICClist[j];
sigmasqb = sigmasqblist[j];
sigmasqe = sigmasqelist[j];
sigmasql = sigmasqllist[j];

theta<-NULL
conflow<-NULL
confupper<-NULL
for (k in 1:mcreps){
  T<-round(runif(sample_length,min=0,max=1))
  Bioleffectvec = rnorm(sample_length,mean=0.5,sd=sqrt(sigmasqb))
  labeffectvec = rnorm(sample_length,mean=0,sd=sqrt(sigmasql))
  erreffect=rnorm(sample_length,mean=0,sd=sqrt(sigmasqe))

  Myresponse<-Bioleffectvec+labeffectvec+erreffect

  lamda_x<-exp(alpha0+beta1*Bioleffectvec+beta2*T+beta3*T*Bioleffectvec)
  lamda_x
  si<-NULL
  for(i in 1:sample_length)
  {
    si[i]<-rexp(n=1,rate=lamda_x[i])
  }
  si  # the si are the survivl times

  #### NO CENSORING YET ###
```

```
   t0<-rep(t04param,sample_length)
   status<-NULL
   status[which(si<=t0)]="D"
   status[which(t0<si)]="C"
   status1<-NULL
   status1[which(status=="C")]=1
   status1[which(status=="D")]=0
   status1

   dataframe<-data.frame(event=status1,trt=T,Y1=Myresponse)
   mytrtsel = trtsel(event="event",trt="trt",marker="Y1",data=dataframe,
               study.design="randomized cohort",link="cloglog",default.trt="trt all")
   #no computation of confident interval
   tmp = eval.trtsel(mytrtsel,bootstraps=0,alpha=0.01)

   theta[k]<-as.numeric(tmp$estimates[8])

 }
 thetanew<-mean(theta)
if (ICC ==1)
{
 setEPS()
 postscript(paste("~/Desktop/ICC/", ICC, ".eps"))
 plot(mytrtsel, bootstraps=100, plot.type="risk", show.marker.axies=TRUE)
 dev.off()
}
```

```
  return(c(ICC,thetanew,t0[1]))
  #return(theta)
}

result1

#hist(result1)

setEPS()
postscript(paste("~/Desktop/ICC_theta_no_bias_", plot_index, ".eps"))
plot(result1[,1],result1[,2],ylim=c(0.95*min(result1[,2]),
    1.05*max(theta, max(result1[,2]))), ylab="estimated theta",
    xlab="ICC",main="estimated theta vs ICC")
abline(h=theta, col="red", lty=2)
abline(a= 0, b=theta,col="red", lty=2)
dev.off()
```

A.1.2 With Censoring Data

```
rm(list=ls())
#######calculate true Theta###########
normalmean = 0.5;
normalvar = 1/12;
mcreps = 2000;
sample_length <- 300
#sample_length <- 1000
T<-round(runif(sample_length,min=0,max=1))
library(survival)
```

```
ICClist = (0:20)/20;

sigmasqblist = ICClist/12;

sigmasqelist = (1/24)*(1-ICClist);

sigmasqllist = sigmasqelist;

Bioleffectvec<-rnorm(1e7,mean=normalmean,sd=sqrt(normalvar))


alpha0 = log(0.6086)

beta1 = -3.145

beta2 = -3.145

beta3 =6.290

plot_index = 1

rate = 0.05


#alpha0 = log(0.2454)

#beta1 = -1.169

#beta2 = -1.169

#beta3 =2.337

#plot_index = 2

#rate = 0.06


#alpha0 = log(0.0262)

#beta1 = 3.145

#beta2 = 3.145

#beta3 =-6.290

#plot_index = 4

#rate = 0.05
```

```
#alpha0 = log(0.1386)
#beta1 = 0
#beta2 = 1.479
#beta3 =-3.145
#plot_index = 5
#rate = 0.055

t0percentile = 0.5;
# This below code sets t0 to be about the approx median of the survival times
#  under control, to change, change t0percentile
t04param = -log(t0percentile)*exp(-(alpha0+beta1*normalmean));
t04param;

# delta_Y = exp(0-t04param*exp(alpha0+beta1*Bioleffectvec+beta2+beta3*Bioleffectvec))
# -exp(-t04param*exp(alpha0+beta1*Bioleffectvec))

delta_Y = 1 - exp(-t04param*exp(alpha0))^exp(beta1*Bioleffectvec) -1 +
exp(-t04param*exp(alpha0))^exp(beta1*Bioleffectvec+beta2+beta3*Bioleffectvec);

stat_delta_Y <-NULL
stat_delta_Y[which(delta_Y<0)]=1
stat_delta_Y[which(delta_Y>=0)]=0
P_neg = sum(stat_delta_Y)/length(delta_Y)

B_neg = 0-sum(delta_Y*stat_delta_Y)/length(which(stat_delta_Y==1))
B_neg
theta <-B_neg*P_neg
```

```
theta
########calculate estimated Theta and plot ICC-Theta##########
ICC<-NULL
#theta_hat_e<-NULL
#theta_hat_mod<-NULL
library(doParallel)
mycluster = makeCluster(3)
registerDoParallel(mycluster)
getDoParWorkers()
library(foreach)
result1 <- foreach(j = 1:length(ICClist),.combine=rbind) %dopar% {
  library(TreatmentSelection)
  ICC = ICClist[j];
  sigmasqb = sigmasqblist[j];
  sigmasqe = sigmasqelist[j];
  sigmasql = sigmasqllist[j];

  theta<-NULL
  conflow<-NULL
  confupper<-NULL
  proportion<-NULL
  for (k in 1:mcreps){
    T<-round(runif(sample_length,min=0,max=1))
    Bioleffectvec = rnorm(sample_length,mean=0.5,sd=sqrt(sigmasqb))
    labeffectvec = rnorm(sample_length,mean=0,sd=sqrt(sigmasql))
    erreffect=rnorm(sample_length,mean=0,sd=sqrt(sigmasqe))
```

```
    Myresponse<-Bioleffectvec+labeffectvec+erreffect

    lamda_x<-exp(alpha0+beta1*Bioleffectvec+beta2*T+beta3*T*Bioleffectvec)
    lamda_x
    si<-NULL
    for(i in 1:sample_length)
    {
        si[i]<-rexp(n=1,rate=lamda_x[i])
    }
    si
    wi<-rexp(n=sample_length,rate=rate)

    ####ci could choose 10, and 20 as threshold)###

    wi
    status<-NULL
    status[which(si<=wi)]="D"
    status[which(wi<si)]="C"

mytime = si;
mytime[which(wi<si)] = wi[which(wi<si)]

    index<-which(status=="D" | wi>= t04param)

   event<- as.numeric(mytime>t04param)
    event<-event[index]
    Myresponse<-Myresponse[index]
```

```
    T<-T[index]
    #status<-status[index]

    dataframe<-data.frame(event=event, trt=T, Y1=Myresponse)
    mytrtsel = trtsel(event="event",trt="trt",marker="Y1",data=dataframe,
              study.design="randomized cohort",link="cloglog",default.trt="trt all")
    #no computation of confident interval
    tmp = eval.trtsel(mytrtsel,bootstraps=0,alpha=0.01)

    proportion[k]<-sum(as.numeric(status=="C"))/length(status)
    #proportion[k]<-sum(event)/length(event)
    #proportion[k]<-(sample_length-length(index))/sample_length
    theta[k]<-as.numeric(tmp$estimates[8])
  }
  thetanew<-mean(theta)
  p<-mean(proportion)

  return(c(ICC, thetanew, t04param, p))
}

result1

setEPS()
postscript(paste("~/Desktop/censor", plot_index, ".eps"))
plot(result1[,1],result1[,2],ylim=c(0.95*min(result1[,2]),
    1.05*max(theta,max(result1[,2]))), ylab="estimated theta",
    xlab="ICC",main="estimated theta vs ICC")
```

```
abline(h=theta, col="red", lty=2)
dev.off()
```

A.2 ICC Estimation from Real Data

```
#Estimated ICC of Exp1A#
library(lme4)
library(forecast)
data<-read.csv('~/Documents/STUDY/UGA/Thesis/Thesis data/Exp1Adata_for_KD.csv')


data1<-NULL
Lab_list<-c('G','B','E','D','H','C','A','F')
for(i in 1:nrow(data))
{
for(j in 1:(length(Lab_list)))
{
ID<-data$ID[i]
Lab<-Lab_list[j]
Myresponse<-as.numeric(data[i,j+1])
data1<-rbind(data1, cbind(ID,Lab,Myresponse))
}
}
data1<-data.frame(data1)
data1$Myresponse<-as.numeric(as.character(data1$Myresponse))
lambda<-BoxCox.lambda(data1$Myresponse)
#get the scatter plot#
data<-BoxCox(data, lambda)
pairs(~G+B+E+D+H+C+A+F, data=data,pch=20)
```

```
data1$Myresponse<-BoxCox(data1$Myresponse,lambda)


lm<-lmer(Myresponse ~ (1|ID)+(1|Lab), data1)
summary(lm)
ICC<-3.4951/(3.4951+0.5830+0.5774)
ICC


#Estimated ICC of Exp1B#
data<-read.csv('~/Documents/STUDY/UGA/Thesis/Thesis data/Exp1Bdata_for_KD.csv')
data1<-NULL
Lab_list<-c('G','B','E','D','H','C','A','F')
for(i in 1:nrow(data))
{
for(j in 1:(length(Lab_list)))
{
ID<-data$ID[i]
Lab<-Lab_list[j]
Myresponse<-as.numeric(data[i,j+1])
data1<-rbind(data1, cbind(ID,Lab,Myresponse))
}
}
data1<-data.frame(data1)
data1$Myresponse<-as.numeric(as.character(data1$Myresponse))
lambda<-BoxCox.lambda(data1$Myresponse)
#get the scatter plot#
data<-BoxCox(data, lambda)
```

```
pairs(~G+B+E+D+H+C+A+F, data=data,pch=20)


data1$Myresponse<-BoxCox(data1$Myresponse,lambda)


lm<-lmer(Myresponse ~ (1|ID)+(1|Lab), data1)
ICC<-2.2415/(2.2415+0.5734+0.8439)
```

APPENDIX B

ADDITIONAL RESULT

B.1 Summary Result for Random Effects Model

B.1.1 Experiment A

```
> summary(lm)
Linear mixed model fit by REML ['lmerMod']
Formula: Myresponse ~ (1 | ID) + (1 | Lab)
Data: data1

REML criterion at convergence: 2253.9

Scaled residuals:
    Min      1Q     Median    3Q     Max
-3.9099 -0.6054   0.0379  0.5510   3.5034

Random effects:
 GroupsName            Variance   Std.Dev.
 ID(Intercept)         3.4951   1.8695
 Lab(Intercept)        0.5830   0.7635
 Residual              0.5774   0.7599
Number of obs: 800, groups:ID, 100; Lab, 8
```

```
Fixed effects:
          Estimate   Std. Error     t value
(Intercept)4.3090       0.3295       13.08
```

B.1.2 Experiment B

```
> summary(lm)
Linear mixed model fit by REML ['lmerMod']
Formula: Myresponse ~ (1 | ID) + (1 | Lab)
Data: data1

REML criterion at convergence: 2479.7

Scaled residuals:
    Min     1Q     Median    3Q    Max
-7.7733 -0.3908   0.0165  0.4821  3.4763

Random effects:
 GroupsName           Variance    Std.Dev.
 ID(Intercept)         2.2415   1.4972
 Lab(Intercept)        0.5734   0.7573
 Residual              0.8493   0.9216
Number of obs: 800, groups:ID, 100; Lab, 8

Fixed effects:
         Estimate   Std. Error     t value
(Intercept)3.6883      0.3085       11.96
```