Testing Equality of Competing Risks of Patient Discontinuation across Multiple Disease States

Eric Nantz
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Recommended Citation
https://scholarworks.wmich.edu/dissertations/689
TESTING EQUALITY OF COMPETING RISKS OF PATIENT DISCONTINUATION ACROSS MULTIPLE DISEASE STATES

by

Eric Nantz

A Dissertation
Submitted to the
Faculty of the Graduate College
in partial fulfillment of the
requirements of the
Degree of Doctor of Philosophy
Department of Statistics
Advisor: Michael R. Stoline, Ph.D.

Western Michigan University
Kalamazoo, Michigan
June 2009
TESTING EQUALITY OF COMPETING RISKS OF PATIENT DISCONTINUATION ACROSS MULTIPLE DISEASE STATES

Eric Nantz, Ph.D.
Western Michigan University, 2009

Previous statistical analyses of patient discontinuation in clinical trials have used discontinuation status as the response of interest. These analyses assume that the risks of discontinuation for specific reasons (lack of efficacy, adverse events, other reasons) are independent of each other and that significant risk factors for patient discontinuation have the same effect on the different causes of discontinuation. However, it is possible that the underlying risks of discontinuation for specific reasons could be related and that risk factors for one type of discontinuation could have a very different effect on another type of discontinuation. The competing risks methodology can be applied to test for significant differences between the different risks of discontinuation and determine significant predictors of discontinuation specific to each type of discontinuation. These competing risks methods are applied to real-world clinical trial data in multiple disease states, which have not been applied previously in the analysis of patient discontinuation.
Copyright by
Eric Nantz
2009
ACKNOWLEDGMENTS

This dissertation marks the end of an incredible journey to obtain this degree, and it would not have been possible without the contributions of my colleagues and family. The chair of my dissertation committee, Dr. Michael Stoline, provided valuable support, encouragement, and guidance from the very beginning to completion of this dissertation. I am so grateful for his patience and willingness to explore this research with enthusiasm. I hope to demonstrate his dedication to his career in my career. I would also like to thank Dr. Joseph McKean for giving me the opportunity to gain valuable experience in statistical consulting and innovative research in statistics, and Dr. Rajib Paul for providing reviews and positive feedback on this dissertation.

I would also like to thank Dr. Hong Liu-Seifert of Eli Lilly & Company, who gave me the opportunity to work with her on the analysis of patient discontinuation. I cannot fully express how much I have learned from her and how grateful I am for her continued support and advice. I would also like to thank Eli Lilly & Company for the use of clinical trial data for this dissertation.

Finally, I am forever grateful to my family. My loving wife Yuqin always made sure I was doing my research and did everything she could help me during the smooth and rough times, and my mother Judi always provided the support and reassurance I needed to meet every challenge and succeed. Lastly, I thank
Acknowledgments—Continued

my grandfather Don, whose intelligence and work ethic provided the best example needed to complete this research and succeed in my career.

Eric Nantz
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ......................................................................................... ii

LIST OF TABLES ................................................................................................. viii

LIST OF FIGURES ................................................................................................. x

CHAPTER

I. BACKGROUND ................................................................................................. 1

1.1 Introduction ............................................................................................... 1

1.2 Literature on Statistical Analysis of Patient Discontinuation .................. 3

1.3 Description of Data .................................................................................... 4

1.4 Application of Competing Risks to Discontinuation Data......................... 8

1.5 Dissertation Outline .................................................................................. 10

II. EXISTING THEORY ......................................................................................... 11

2.1 Introduction ................................................................................................ 11

2.2 Basic Concepts of Survival Analysis .......................................................... 11

2.2.1 Introduction .......................................................................................... 11

2.2.2 Parameters of Interest ......................................................................... 12

2.2.3 Estimation of Parameters ...................................................................... 14

2.3 Basic Concepts in Competing Risks ............................................................ 16

2.3.1 Bivariate Approach .............................................................................. 16
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.2 Latent Failure Time Approach</td>
<td>20</td>
</tr>
<tr>
<td>2.3.3 Discussion of Approaches to Competing Risks</td>
<td>21</td>
</tr>
<tr>
<td>2.3.4 Estimation of Parameters</td>
<td>22</td>
</tr>
<tr>
<td>2.3.5 Modeling in Competing Risks</td>
<td>24</td>
</tr>
<tr>
<td>2.4 Counting Process and Martingale Methodology</td>
<td>27</td>
</tr>
<tr>
<td>2.4.1 Martingale Preliminary Concepts</td>
<td>28</td>
</tr>
<tr>
<td>2.4.2 Counting Process Preliminary Concepts</td>
<td>33</td>
</tr>
<tr>
<td>2.4.3 Application to Survival Analysis</td>
<td>38</td>
</tr>
<tr>
<td>2.4.4 Application to Competing Risks Analysis</td>
<td>41</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>III. INFERENCE PROCEDURES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>44</td>
</tr>
<tr>
<td>3.2 Martingale Central Limit Theorem</td>
<td>44</td>
</tr>
<tr>
<td>3.3 KLY Test Procedure Comparing Competing Risks</td>
<td>48</td>
</tr>
<tr>
<td>3.4 TKC Tests</td>
<td>54</td>
</tr>
<tr>
<td>3.4.1 Introduction</td>
<td>54</td>
</tr>
<tr>
<td>3.4.2 Asymptotic Distributions</td>
<td>58</td>
</tr>
<tr>
<td>3.4.3 Weight Functions</td>
<td>65</td>
</tr>
</tbody>
</table>
CHAPTER

IV. MONTE CARLO SIMULATION STUDY ................................. 67
   4.1 Simulation Procedures ........................................... 67
   4.2 Simulation Results ............................................... 72
       4.2.1 Introduction ............................................... 72
       4.2.2 Impact of Censoring Results ............................. 73
       4.2.3 Impact of Weight Functions Results .................... 73

V. APPLICATION OF COMPETING RISKS METHODS .................... 76
   5.1 Introduction ..................................................... 76
   5.2 Baseline Results ................................................ 76
   5.3 Competing Risks Testing Results .............................. 78
   5.4 Regression Modeling Results ................................... 88

VI. SUMMARY AND CONCLUSIONS ........................................ 95
   6.1 Summary ......................................................... 95
   6.2 Conclusions ..................................................... 97
   6.3 Additional Research ........................................... 97

REFERENCES .............................................................. 99

APPENDIX ................................................................. 106
   A. PROGRAM CODE .................................................. 106
       A.1 Simulation Driver Program ................................ 106
Table of Contents—Continued

APPENDIX

A.2 Data Simulator .............................................. 111
A.3 Test Simulator .............................................. 112
A.4 KLY Test Procedure Program ................................. 113
A.5 TKC Test Procedure Program ................................. 116
A.6 GAD Treatment Assignment Plot Generator ................. 129
A.7 Fine and Gray Regression Model Selection Program ......... 139
A.8 Fine and Gray Regression Model Selection Runner .......... 143
A.9 Cox Regression Model Selection Program ..................... 147
A.10 Cox Regression Model Selection Runner ..................... 151
LIST OF TABLES

1. Configurations of TKC Type I Error Rate Simulations ......................... 71
2. Configurations of TKC Power Simulations ........................................ 72
3. Empirical Type I Error Rates Under Different Censoring Conditions .......... 74
4. Empirical Power for TKC Tests ....................................................... 75
5. Number (percent) of Patients who DC or Complete ............................. 77
6. Descriptive Statistics by Disease State ............................................. 77
7. TKC Test 3 and KLY Test Results .................................................... 79
8. TKC Test Statistics for Risks of DC with Treatment Assignment ............. 80
9. TKC Test 2 Partitions for Risks of DC with Treatment Assignment .......... 81
10. TKC Test 3 Partitions for Risks of DC with Treatment Assignment ......... 83
11. TKC Test 3 Partitions for Risks of DC (DC-OTH Censored) .................. 83
12. TKC Test Statistics and P-values for Risks of DC with Study Location ...... 85
13. TKC Test 2 Partitions for Risks of DC with Study Location .................. 86
14. TKC Test 3 Partitions for Risks of DC with Study Location .................. 88
15. CRR Model Selection Results for DC due to AE ................................. 91
16. Cox PH Model Selection Results for DC due to AE ............................ 91
17. CRR Model Selection Results for DC due to LOE ............................... 92
List of Tables—Continued

18. Cox PH Model Selection Results for DC due to LOE .................. 92
19. CRR Model Selection Results for DC due to OTH .................. 93
20. Cox PH Model Selection Results for DC due to OTH .................. 93
LIST OF FIGURES

1. Cumulative Incidence Curves ........................................ 78
2. CCSH Curves for DC Categories and Treatment Groups .......... 80
3. CCSH Curves for GAD and Treatment Groups ...................... 82
4. CCSH Curves for MDD and Treatment Groups ..................... 82
5. CCSH Curves for GAD and Treatment Groups ..................... 84
6. CCSH Curves for MDD and Treatment Groups ..................... 84
7. CCSH Curves for DC Categories and Study Location ............ 86
8. CCSH Curves for GAD and Study Location ......................... 87
9. CCSH Curves for MDD and Study Location ......................... 87
10. CCSH Curves for GAD and United States .......................... 88
11. CCSH Curves for MDD and United States .......................... 89
CHAPTER I

BACKGROUND

1.1 Introduction

The goal of a clinical trial is to determine if a treatment is effective and if it is safe. According to Meinert (1986), a clinical trial is defined as “a planned experiment designed to assess the efficacy of a treatment in humans by comparing outcomes in a group of patients treated with a test treatment with those observed in a comparable group of patients receiving a control treatment, where patients in both groups are enrolled, treated, and followed over the same time period.” Implementing randomization and double-blind procedures help to minimize the possible treatment effect bias. Leon et al (2007) define bias as the difference between the unknown population treatment effect and estimate of the treatment effect formed using the sample data. Randomization helps minimize bias by ensuring that each patient has an equal chance of being assigned to a treatment group. Double-blind in a clinical trial helps to minimize bias by eliminating the possible preconceived opinions about the treatment had either the patient or the investigator known which treatment the patient received. These are two features of a randomized controlled clinical trial (RCT).

Patients are given information about the study before they decide to vol-
unteer for the study, known as informed consent. Part of the informed consent process is the right for the patient to discontinue from the study at any time. Ideally, all of the patients that enroll for the clinical trial will be able to continue to the conclusion of the trial. However, in clinical trials with human subjects, patient discontinuation from the trial is almost unavoidable. It is recommended that the flow of participants through each stage of the study, including patient attrition rates be documented in the publication of the results of the RCT (Altman et al., 2001).

The loss of patients in a clinical trial can have a wide range of impact on the validity of results. The fact that a patient decides to discontinue from the study can be viewed as a form of self-selection, which in turn could compromise the randomization principle (Leon et al., 2007). If there is a treatment effect present on the outcome of interest, then a reduced number of patients in the study will reduce the power of the trial for detecting a significant treatment effect (Siddiqi, 2006). A reduction in the sample size for a clinical trial can threaten the internal validity of results and limit the generalizability of the results if there are traits of the patients who discontinued from the study that are not the same as the traits of patients who completed the study (Leon et al., 2007). These consequences of patient attrition highlight the importance of understanding the processes and risk factors influencing patient attrition. Results from statistical analysis techniques aimed at determining significant differences between these groups of patients and
their risk factors could help improve the quality of the RCT.

1.2 Literature on Statistical Analysis of Patient Discontinuation

The consequences of attrition in clinical interventions have motivated research on investigating distinguishing characteristics and drivers of patient attrition. Much of the previous research examines discontinuation for any reason as the outcome of interest. In general, a variable set of significant predictors have been determined using established statistical modeling and inferential techniques. Logistic regression models have found early symptom response in schizophrenia clinical trials (Liu-Seifert et al., 2005), age, education, and perceived mental health in Major Depressive Disorder patients (Warden et al., 2007); caregiver age in pediatric asthma intervention (Zebracki et al., 2003); age, ethnicity, and perceived severity of illness in HIV intervention trials (DiFranceisco, 1998) predicted probability of treatment discontinuation. Cox proportional hazard models have found that baseline and change in response from previous visit in schizophrenia clinical trials (Rabinowitz & Davidov, 2008) were significant risk factors of study discontinuation. Nantz et al (2009) determined that both early and visit-wise lack of therapeutic response and adverse event severity were significant risk factors of discontinuation in multiple disease states through logistic and Cox regression models. Generalized estimating equations (GEE) models have found that patient-physician discordance in clinical interventions for Fibromyalgia patients (Sewitch
et al., 2004) and severity of depression, duration, and education in cancer patients treated with chemotherapy (Siddiqi, 2006) were significant characteristics of attrition. Given the variable set of characteristics predicting patient attrition among patients treated for different disease types, it is reasonable to assume that the nature of the disease is a contributing factor to this variation.

The results of analyses of patient discontinuation can play a key role in the statistical analysis of therapeutic response in clinical interventions. Selection models, first proposed by Heckman (1976), incorporate predicted probabilities of dropout for each subject as covariates in a longitudinal data model for therapeutic response. Another class of models called pattern-mixture models proposed by Little (1993) divides subjects into groups based on their missing data pattern and these groups of subjects are then analyzed separately. See Hedeker and Gibbons (1997) for an application of pattern-mixture models to the analysis of schizophrenia patients.

1.3 Description of Data

The data sources for this research are twenty-two randomized, double-blind, placebo-controlled clinical trials within the Eli Lilly & Company database assessing the safety and efficacy of duloxetine (brand name Cymbalta), a serotonin-norepinephrine reuptake inhibitor. These clinical trials involve the treatment of the following disease states: diabetic neuropathic nerve pain (DPNP), fibromyal-
gia (FM), generalized anxiety disorder (GAD), and major depressive disorder (MDD). Analyses are conducted separately for each disease state due to the substantial clinical differences between the patient populations.

All clinical trials within a disease state use the same primary efficacy measurements. The primary efficacy measurements for each disease state are the following:

- DPNP: Weekly mean of the 24-hour average pain rating
- FM: 24-hour average pain item from the Brief Pain Inventory (BPI)
- GAD: Hamilton Anxiety Rating Scale Total Score (HAMA)
- MDD: Total score of the 17-item Hamilton Depression Rating Scale (HAMD17)

Each adverse event's severity was assessed using a severity scale with the following values:

- Severity Rating = 1: Mild adverse event (does not disrupt patient’s usual activity)
- Severity Rating = 2: Moderate adverse event (mild disruption in patient’s usual activity)
- Severity Rating = 3: Severe adverse event (major disruption in patient’s usual activity)
The clinical trials used in this analysis all begin with a screening phase, in which the eligibility status of patients is determined. The patients that pass the screening phase proceed to the acute therapy phase, where patients are randomized to be treated with placebo, duloxetine, or in some studies a comparative treatment. If a patient completes the acute therapy phase, they enter a posttherapy drug tapering phase. Some studies include an open-label extension phase in which all patients completing the acute therapy phase are randomly assigned to different dosages of duloxetine treatment. This dissertation focuses on the acute therapy phase of these clinical trials.

When a patient discontinues from a clinical trial, their investigator are required to record the date and reason for discontinuing the trial. The reasons for discontinuation consist of the following:

- Lack of Efficacy (LOE): The patient decides that their symptom improvement was inadequate.

- Adverse Event (AE): The patient experiences an unwanted side effect, and the investigator specifies the event.

- Subject Decision: The patient wishes to discontinue treatment for personal reasons.

- Lost to Follow Up: The patient misses a scheduled visit and the investigator cannot reach them using phone or mail.
- Physician Decision: The physician decides the patient should be discontinued from the treatment for reasons other than LOE or AE.

- Protocol Violation of Entry Criteria: The patient was mistakenly enrolled based on specific entry criteria, or the patient did not follow the procedures and requirements specified by the protocol.

- Sponsor Decision: The sponsor decided the patient should be discontinued after consulting with the patient's investigator.

The previous research on patient attrition has combined each different category of discontinuation into one event type: discontinuation or completion. In this research, there are four event types of interest:

1. Discontinuation due to AE
2. Discontinuation due to LOE
3. Discontinuation due to other reasons (consisting of subject decision, lost to follow up, physician decision, protocol violation of entry criteria, and sponsor decision)
4. Completion of acute therapy phase

The competing risks approach to the analysis of these data will treat the events of DC due to LOE, DC due to AE, and DC due to other reasons as separate end points, while censoring the patients that complete the acute therapy phase.
1.4 Application of Competing Risks to RCT Discontinuation Data

In previous research, the time to patient discontinuation has been analyzed using standard survival analysis techniques. In this approach, discontinuation due to any reason is the event of interest, with censoring occurring if the patient completes the treatment. One of the assumptions of this method is that the reasons for patient discontinuation have the same level of importance. This may not be the case for a particular clinical trial. For instance, patient discontinuations due to lack of efficacy or adverse events may have different consequences than patient discontinuation due to "other reasons". Also, it is possible that a treatment related factor of interest such as change in efficacy could have different effects for the different reasons of discontinuation. Within the sample of patients in a clinical trial, the risk of discontinuation due to lack of efficacy may act significantly different than the risk of discontinuation due to adverse events.

Due to the multiple categories of discontinuation, the competing risk approach is a natural choice for the method of analysis. The competing risk approach accounts for possible dependencies or relationships between the competing risks in the estimation of cumulative incidence, which is the probability that a subject will experience a particular event by time $t$. Traditional survival methods such as the Kaplan-Meier estimation of survival and distribution functions and the Cox proportional hazards regression model for identifying significant covariates have analogous techniques in the competing risk situation. Use of the Kaplan-Meier
method for estimation of cumulative incidence for event types is incorrect in the situation of competing risks because it censors subjects that experience a different event type than the event of interest. Also, traditional survival methods cannot assess the differences between competing risks within one sample of subjects. In the competing risks framework, the cumulative incidence functions for competing risks can be directly compared within the same sample using procedures such as the KLY test procedure by Kochar, Lam, and Yip (Kochar et al., 2002) and the TKC test procedures developed by Tiwari et al (Tiwari et al., 2006).

Martingale and counting process theory, originally developed for event history analysis, has emerged as an appealing technique for deriving theoretical results and modeling strategies in survival analysis, due to the fact that recording events lends naturally to using counting processes. Survival and competing risks data are natural applications of counting process and martingale theory. The main advantages of applying martingale and counting process theory to survival and competing risks analyses are intuitive derivations for asymptotic results, unified development of inference techniques.

The goal of this dissertation is to apply the competing risks methodologies with martingale and counting process theory as the driver of theoretical results to the analysis of patient discontinuation data. Estimation of cumulative incidences of discontinuation, testing risks of discontinuation among patients within clinical trials, and modelling of the risk of discontinuation with significant risk factors will
be performed.

1.5 Dissertation Outline

The dissertation is organized as follows. Chapter II describes the parameters and estimation for both traditional survival analysis and competing risks analysis, along with a survey of the counting process and martingale methodology. Chapter III introduces the martingale central limit theorem and outlines the KLY and TKC test procedures for comparing competing risks. In chapter IV, a comprehensive simulation study is conducted to examine the performance of the TKC test procedures in terms of type I error and power under different censoring conditions and departures from the null hypotheses. Clinical trial data from two disease states are used to illustrate the application of the competing risks techniques and inference procedures in chapter V. Finally, a summary of the results along with conclusions and additional research possibilities in chapter VI concludes the dissertation.
CHAPTER II

EXISTING THEORY

2.1 Introduction

This chapter contains a review of the statistical theory and methodologies used in this research. A review of standard survival analysis theory is discussed in section 2.2. Theory and methodologies of competing risks are discussed in section 2.3. A discussion of counting process and martingale theory is contained in section 2.4.

2.2 Basic Concepts of Survival Analysis

2.2.1 Introduction

Survival analysis is a set of statistical techniques and methodologies that are tailored to data involving the time until the occurrence of an event of interest. Survival analysis has been applied to wide variety of fields, from medical research to reliability of machines. An important characteristic of survival data is the fact that for some subjects the event of interest has not occurred at the end of the experiment's observation time, or that the subject has been lost for reasons unrelated to the event. This concept is known as censoring. In partic-
ular, right censoring occurs when the subject did not have the event of interest during the time that the subject was participating in the study. For detailed discussions of the other types of censoring, see Kalbfleisch and Prentice (Kalbfleisch & Prentice, 2002), Klein and Moeschberger (Klein & Moeschberger, 2005), and Lawless (Lawless, 2002). In the context of this dissertation, the event of interest is discontinuation from the study.

2.2.2 Parameters of Interest

Let $T$ denote the continuous random variable representing the time until discontinuation from the study. Let $\delta$ represent the censoring variable, in which $\delta = 1$ means the patient experienced the event of interest during the study period, and $\delta = 0$ if the patient did not discontinue during the study period. The survival random variable $T$ has a cumulative distribution function $F(t)$ and a probability density function $f(t)$. The cumulative distribution function of $T$ represents the probability that the amount of time a subject remains in the experiment is less than or equal to some specific time $t$, and is given by

$$F(t) = P(T \leq t) = \int_0^t f(x)\,dx$$  \hspace{1cm} (2.1)

One of the functions that draws the highest interest from researchers is the survival function, denoted by $S(t)$, which describes the probability that a subject
survives longer than a specific time $t$:

$$S(t) = P(T > t) = 1 - F(t) \quad (2.2)$$

The hazard function $h(t)$, otherwise known as the conditional failure rate, expresses the probability of the subject experiencing the event of interest during a very small interval of time, conditional on the subject staying event-free until time $t$:

$$h(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt \mid T > t)}{dt} \quad (2.3)$$

The cumulative hazard function is defined by:

$$H(t) = \int_0^t h(x)dx \quad (2.4)$$

Relationships among these quantities are useful for estimation:

$$h(t) = \frac{f(t)}{S(t)}, \quad (2.5)$$

$$h(t) = -\frac{d}{dt} \log S(t), \text{ and} \quad (2.6)$$

$$S(t) = e^{-H(t)} = e^{-\int_0^t h(x)dx} \quad (2.7)$$
2.2.3 Estimation of Parameters

The survival function $S(t)$ is commonly estimated using the Kaplan-Meier product-limit estimator originally developed by Kaplan (Kaplan & Meier, 1958). Let $n$ denote the number of subjects with censored and uncensored event times $t_1, t_2, \ldots, t_n$. Let $0 < t_{(1)} < t_{(2)} < \cdots < t_{(r)}$ be the $r$ ordered distinct uncensored event times. Let $n_i$ denote the number of subjects remaining in the trial just before time $t_{(l)}$. Let $d_l$ denote the number of subjects who experience the event at time $t_{(l)}$. A critical assumption made in the derivation of the Kaplan-Meier estimator is that the censoring process is non-informative, meaning that the risk for being censored at time $t$ does not depend on the risk for experiencing an event at time $t$. The probability of surviving through the $l$th constructed time interval from $t_{(l)}$ to $t_{(l+1)}$ for $l = 1, ..., r$ (where $t_{(r+1)}$ is defined to be $\infty$) and all preceding intervals is estimated using the Kaplan-Meier estimator defined as

$$
\hat{S}(t) = \prod_{j=1}^{l} \left(1 - \frac{d_j}{n_j}\right).
$$

Estimating the hazard function has proven to be difficult in the literature. However, the cumulative hazard function can be estimated using the Nelson-Aalen estimator. Originally proposed by Nelson in quality improvement settings (Nelson, 1969),(Nelson, 1972), and later extended to event history analysis by Aalen (Aalen, 1975),(Aalen, 1978), the Nelson-Aalen estimator of the cumulative hazard function
$H(t)$ is defined as

$$\hat{A}(t) = \sum_{j=1}^{t} \frac{d_{j}}{n_{j}} \quad (2.9)$$

Both the Kaplan-Meier and Nelson-Aalen estimators do not take into account possible covariates that may influence the survival and hazard functions. A statistical model called the Cox proportional hazards (PH) model (Cox, 1972) provides a way to incorporate covariate effects in the modeling of the hazard function. Let $x = (x_1, \ldots, x_p)'$ denote the $p \times 1$ vector of covariates and $\beta = (\beta_1, \ldots, \beta_p)'$ denote the $p \times 1$ vector of unknown regression coefficients. The proportional hazards model is the following:

$$h(t \mid x) = h_0(t)e^{\beta'x} \quad (2.10)$$

where $h(t \mid x)$ is the hazard function at time $t$ with time-independent covariate vector $x$, $h_0(t)$ is the baseline hazard, and $\beta$ is the coefficient vector for $X(t)$. Estimates of the $\beta$-coefficients are found by maximizing the likelihood function

$$L(\beta) = \prod_{j=1}^{t} \frac{\exp(\beta'x_{(j)})}{\sum_{i \in R(t_{(j)})} \exp(\beta'x_i)} \quad (2.11)$$

where $x_{(j)}$ is the vector of covariates for the subject who experiences the event at the $j$th ordered event time $t_{(j)}$ and the summation in the denominator is the sum of the values of $\exp(\beta'x)$ over all subjects who are at risk at time $t_{(j)}$.

In the situation of covariates that depend on time, such as visit-wise mea-
measurements of efficacy, another version of the Cox regression model is the following:

\[ h(t \mid \mathbf{x}(t)) = h_0(t)e^{\mathbf{x}(t)\mathbf{\beta}} \]  

(2.12)

where \( \mathbf{x}(t) = (x_1(t), \ldots, x_p(t))' \) is the vector of covariates that may or may not depend on time. The likelihood function used to derive estimates of the \( \beta \)-coefficients is the following:

\[ L(\beta) = \prod_{j=1}^l \frac{\exp(\beta' \mathbf{x}_{(j)}(t_{(j)}))}{\sum_{i \in R(t_{(j)})} \exp(\beta' \mathbf{x}_{(i)}(t_{(i)}))} \]  

(2.13)

2.3 Basic Concepts in Competing Risks

The history of competing risks began in 1760 when Daniel Bernoulli analyzed the risks of death from smallpox compared to death from other causes (David & Moeschberger, 1978). Bernoulli’s work laid the foundation for the theory of competing risks used today. The following definitions and results can be found in Pintilie (Pintilie, 2006) and Crowder (Crowder, 2001). Competing risks is the situation when a subject can experience more than one type of event and the occurrence of one type hinders or prevents the occurrence of other types of events.

2.3.1 Bivariate Approach

The two approaches to describing competing risks are the bivariate approach and the latent failure time approach. The bivariate approach will be
discussed first. In competing risks, the observed data consist of $T$, the time to event or censoring, and $\delta$, the type of event. Assume that there are $p$ different events of interest. If the observation is censored, then $\delta = 0$. If the observation is not censored, then $\delta = j$, where $j > 0$ is the type of event observed. The event time $T$ is a continuous random variable and $\delta$ is a discrete random variable. The following quantities are based on the joint distribution of $T$ and $\delta$. The joint distribution of $T$ and $\delta$ can be specified using two functions called the cumulative incidence function (CIF) and the sub-survivor function.

The CIF, or subdistribution function, is the probability that an event of type $j$ occurs at or before time $t$ while the individual is at risk. The CIF for an event of type $j$ where $j = 1, 2, \ldots, p$ is defined by the joint probability

$$F_j(t) = P(\delta = j \text{ and } T \leq t) \quad (2.14)$$

and the sub-density function is defined as

$$f_j(t) = \frac{\partial F(j, t)}{\partial t} \quad (2.15)$$

The sub-survivor function is the probability that an event of type $j$ does not occur by time $t$ and is represented by the joint probability

$$S_j(t) = P(\delta = j \text{ and } T > t) \quad (2.16)$$
The marginal distribution of $\delta$ is denoted by $p_j = P(\delta = j)$. The CIF is referred to as a subdistribution because at $t = \infty$, $F_j(\infty) = p_j$.

Proof.

\[
F_j(\infty) = P(\delta = j, 0 < t < \infty) = \int_0^{\infty} f(j, t)dt = P(\delta = j)
\]

Hence the CIF is not a proper distribution. It can also be shown that $S_j(0) = p_j$. Also, for a particular event of type $j$, $F_j(t) + S_j(t) = p_j$.

Note that the overall distribution function (the probability that an event of any type occurs at or before time $t$) denoted by $F(t)$, is equal to the sum of the cumulative incidence functions for each event type:

\[
F(t) = P(T \leq t) = \sum_{j=1}^{p} P(T \leq t, \delta = j) = \sum_{j=1}^{p} F_j(t)
\]

(2.17)

where $p$ is the number of possible events. The *cause-specific hazard function* (CSH) for event type $j$ expresses the probability of the subject experiencing the event of type $j$ during a very small interval of time, conditional on the subject staying event-free until time $t$: 
\[ h_j(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt, \delta = j \mid T \geq t)}{dt} \quad (2.18) \]

with the cumulative cause-specific hazard for event type \( j \) defined as

\[ H_j(t) = \int_0^t h_j(s) \, ds. \quad (2.19) \]

The overall hazard function from all events is the same as the hazard function in typical survival data analysis, and it is the sum of all of the cause-specific hazard functions:

\[ h(t) = \sum_{j=1}^{p} h_j(t) \quad (2.20) \]

The relationship between the cause-specific hazard function and the CIF for event type \( j \) is characterized by

\[ F_j(t) = \int_0^t h_j(s) S(s) \, ds \quad (2.21) \]

Another key parameter of interest is a special version of the hazard function called the hazard of the subdistribution (HOS) developed by Gray (Gray, 1988). The HOS for event type \( j \) is explicitly defined by:

\[ \gamma_j(t) = \lim_{dt \to 0} \frac{P(t \leq T < t + dt, \delta = j \mid T \geq t \cup (T \leq t \cap \delta \neq j))}{dt} \quad (2.22) \]
where \( T \geq t \cup (T \leq t \cap \delta \neq j) \) means that either the event of interest did not occur until time \( t \), or that the observation for a subject in the study has stopped because the subject experienced a competing risk event. The relationship between \( \gamma_j(t) \) and the subdensity and subdistribution functions given in 2.15 and 2.17 respectively is given by

\[
\gamma_j(t) = \frac{f_j(t)}{1 - F_j(t)}. \tag{2.23}
\]

### 2.3.2 Latent Failure Time Approach

The other approach for describing competing risks is the latent failure time approach. Let \( T_1, T_2, \ldots, T_p \) denote the latent or unobserved event times for the \( p \) competing risks. Define \( T_j \) as the failure time due to event type \( j \). Since only the first event is observed for each subject, the failure time random variable \( T = \min \{T_1, T_2, \ldots, T_p\} \). The censoring variable \( \delta = 0 \) when the observation is censored, and \( j > 0 \) when the \( j \)th event is observed for \( j = 1, \ldots, p \) possible events. Define a multivariate joint survival function as

\[
S(t_1, \ldots, t_p) = P(T_1 > t_1, T_2 > t_2, \ldots, T_p > t_p). \tag{2.24}
\]

The parameters of interest discussed in the bivariate approach can be expressed in terms of the joint survival function 2.24. The marginal survival function for
event type \( j \) is defined as

\[
S_j(t) = S(t_1 = 0, t_2 = 0, \ldots, t_j = t, \ldots, t_p = 0)
\]  

(2.25)

The subdensity function for event type \( j \) is defined as

\[
f_j(t) = \left( -\frac{\partial S(t_1, t_2, \ldots, t_p)}{\partial t_j} \right)
\]

(2.26)

and the CIF or subdistribution function is defined as

\[
F_j(t) = \int_0^t f(j, s)ds
\]

(2.27)

equivalent to the CIF given in 2.17. The CSH function for event type \( j \) is defined as

\[
h_j(t) = \left. \frac{-\partial \log S(t_1, \ldots, t_p)}{\partial t_j} \right|_{t_1 = t_2 = \ldots = t_p = t}
\]

(2.28)

equivalent to the CSH defined in 2.18.

2.3.3 Discussion of Approaches to Competing Risks

Both the bivariate and latent failure time approaches have been analyzed in the literature on competing risks data. While the latent failure time was the first approach utilized in competing risks models, it also carried a serious issue of non-identifiability. Specifically, unless the marginal survival functions are assumed to
be independent, more than one joint survival distribution can be defined based on
the same marginal survival functions. Cox (Cox, 1959) noted this issue and Tsi-
atis (Tsiatis, 1975) found that different joint distributions can produce the same
marginal distributions without the assumption of independence. With only the
first failure time being observed, the independence assumption for the marginal
survival functions cannot be tested. However, the issue of non-identifiability does
not occur if quantities like the sub-survivor function (2.16) and CSH (2.18) de-
 fined with the bivariate approach are used in modelling competing risks data
(Dewan et al., 2004). For these reasons, the bivariate approach will be used in all
subsequent analyses.

2.3.4 Estimation of Parameters

The following is a heuristic estimation procedure for the CIF function for
event type \( j \). Theoretical details of this estimator are justified using the counting
process and martingale theory given section 2.4. Let \( n \) denote the number of
subjects with censored and uncensored event times \( t_1, t_2, \ldots, t_n \). Let \( 0 < t_{(1)} <
t_{(2)} < \cdots < t_{(r)} \) be the \( r \) ordered distinct event times. \( F_j(t) \) can be calculated as
the summation over all time points \( t_i \leq t \) of the probabilities of failure due to event
type \( j \) at time \( t_i \), given that the subject remained event-free immediately prior to
time \( t_i \). This summation is the Nelson-Aalen estimate for the cumulative CSH for
event type \( j \). Let \( d_{ji} \) denote the number of subjects who failed due to event type \( j \)
at time point $t_i$ and let $n_i$ denote the number of subjects at risk for experiencing an event at time $t_i$. The Nelson-Aalen estimate $\hat{H}_j(t)$ of the cumulative CSH for event type $j$ is calculated as

$$\hat{H}_j(t) = \sum_{i'=1}^l \frac{d_{ji'}}{n_{i'}}$$

Hence the estimate $\hat{F}_j(t)$ for the CIF of event type $j$ is calculated as

$$\hat{F}_j(t) = \sum_{i'=1}^l P(T = t_{i'}, \delta = j)$$

$$= \sum_{i'=1}^l P(T = t_{i'}, \delta = j|T > t_{i'-1}) P(T > t_{i'-1})$$

$$= \sum_{i'=1}^l \hat{H}_j(t_{i'}) \hat{S}(t_{i'-1}).$$

In the testing and modeling procedures to follow, groups of subjects will be analyzed as well. Let $n_k$ denote the number of subjects in group $k$, $k = 1, \ldots, K$. Let $H_{jk}(t)$ and $F_{jk}(t)$ denote the cumulative CSH and CIF for event type $j$ and group $k$. Using the same ordered distinct event times as previously, let $d_{jklt}$ denote the number of subjects who failed due to event type $j$ in group $k$ at time point $t_l$ and let $n_{kl}$ denote the number of subjects in group $k$ at risk for experiencing an event at time $t_l$. The Nelson-Aalen estimate $\hat{H}_{jk}(t)$ of the cumulative CSH for
event type $j$ in group $k$ is calculated as

$$
\hat{H}_{jk}(t) = \sum_{l'=1}^{l} \frac{d_{jkl'}}{n_{kl'}}
$$

and the estimate of the CIF for event type $j$ and group $k$ is calculated as

$$
\hat{F}_{jk}(t) = \sum_{l'=1}^{l} \frac{\hat{H}_{jk}(t_{l'})}{S(t_{l'-1})}.
$$

### 2.3.5 Modeling in Competing Risks

When it is desired to include the effects of covariates, regression models in the situation of competing risks can be used. There are two different regression models that can be applied in the competing risks situation. The first is the Cox PH model applied to a cause-specific hazard $h_j(t)$. Let $x(t) = (x_1(t), \ldots, x_p(t))'$ denote the vector of covariates that may or may not depend on time, and $\beta_j = (\beta_{j1}, \ldots, \beta_{jp})'$ denote the $p \times 1$ vector of unknown regression coefficients for event type $j$. The Cox model applied to the CSH of event type $j$ is given by:

$$
h_j(t \mid x(t)) = h_{j,0}(t)e^{\beta_j'x(t)} \quad (2.29)
$$

where $h_j(t \mid x(t))$ is the cause-specific hazard function corresponding to event type $j$ at time $t$ with possibly time-dependent covariate vector $X(t)$, $h_{j,0}(t)$ is the baseline cause-specific hazard corresponding to event type $j$, and $\beta_j$ is the
coefficient vector for $X(t)$ of event type $j$. In this model, any subject that does not experience any event or experiences a different event than type $j$ are censored.

It has been noted in the literature that interpretation of covariate effects in terms of the subsurvivor function for event type $j$ is very difficult when using the Cox PH model applied to the CSH for event type $j$. The subsurvivor function and CIF are functions of all CSH functions, as seen in equation 2.20, and it is very possible that a covariate could have different effects on the different CSH's. In the survival analysis framework with only one event of interest, interpretations of covariate effects from the Cox PH model can be made in terms of the survival function $S(t)$ as well as the distribution function $F(t)$ since there is a direct one-to-one relationship between the hazard function and the distribution function.

Fine & Gray (Fine & Gray, 1999) developed a proportional hazards model based on the CIF function, in which they utilized the hazard of the subdistribution (HOS) defined in 2.22. Analogous to how the hazard function is related to the survival function given in 2.6,

$$
\gamma_j(t) = -\frac{d}{dt} \log S_j(t)
$$

$$
= -\frac{d}{dt} \log \{1 - F_j(t)\}.
$$

Modeling the cumulative incidence function is accomplished with the Fine & Gray (Fine & Gray, 1999) HOS regression model. Let $x = (x_1, \ldots, x_p)'$ denote the $p \times 1$ vector of covariates and $\beta = (\beta_1, \ldots, \beta_p)'$ denote the $p \times 1$ vector of
unknown regression coefficients. The HOS regression model is given by:

\[ \gamma_j(t \mid x) = \gamma_{j,0}(t)e^{\beta'x} \]  

(2.30)

where \( \gamma_j(t, x) \) is the hazard of the subdistribution corresponding to event type \( j \) at time \( t \) with covariate vector \( x \), \( \gamma_{j,0} \) is the baseline hazard of the subdistribution corresponding to event type \( j \), and \( \beta \) is the coefficient vector for \( x \).

It is important to consider both the Cox PH model applied to CSH's and the HOS regression models in the presence of competing risks. These two regression models help to answer different research questions. The regression model for a cause-specific hazard function can determine the pure effect of covariates on the specific cause of interest, ignoring the occurrences of competing events. The regression model for hazard of the subdistribution can determine the significant predictors of the probability of the event of interest, taking into account the possibility of competing events. Since cumulative incidence is a function of the cause specific hazards for all of the possible events, the effect of a particular covariate on the CIF one type of event depends on both the effect of the covariate on the cause-specific hazard function for the particular event type and the effect on the cause-specific hazard functions for the remaining causes. In contrast, the regression model applied to a particular cause-specific hazard function censors the occurrences of the other competing risk events. Unlike the Cox regression model of the hazard for the event of interest in the typical survival analysis situation,
the Cox regression model applied to the cause-specific hazard function does not model the marginal hazard in the competing risks situation.

2.4 Counting Process and Martingale Methodology

The counting process and martingale methodology has helped establish theoretical results in applications such as survival analysis, competing risks analysis, and recurrent events analysis. The methodology was first used for event history analysis, in which a collection of individuals is observed over time as they transition from one state to another through events, and the probabilities of these events are modeled. A counting process handles recording the events and an intensity process handles the instantaneous probabilities. As a result, both traditional survival and competing risks analyses are two applications of event history analysis. Asymptotic results for survival analysis have been derived using counting process theory, and time-varying covariates are easily incorporated, unlike traditional survival methods. O.O. Aalen's Ph.D. dissertation (1975) marked the origin of applying counting process and martingale theory to survival analysis, and Aalen (1976) also did the same for the competing risks situation. More details on the history of counting process theory can be found in Andersen et al (1993). The following section is a summary of the mathematical details of counting process and martingale theory applicable to survival and competing risk analyses, and will be presented in a similar style as Aalen et al (2008). For a thorough and mathe-
matically detailed discussion, see Fleming & Harrington (1991) and Andersen et al (1993)

2.4.1 Martingale Preliminary Concepts

The martingale methodology relies on stochastic or random processes and probability theory. First we introduce some basic definitions used in the theoretical results.

Definition 2.4.1.

(i) Let \( \Omega \) be the sample space of all outcomes. A \( \sigma \)-algebra is a non-empty collection of subsets, called events, of \( \Omega \) that consists of the sample space \( \Omega \), the empty set \( \emptyset \), the union of those events, the intersections of those events, and the compliments of those events.

(ii) A history \( \{ \mathcal{F}_t : t \geq 0 \} \) is an increasing sequence of sub-\( \sigma \)-algebras from \( \Omega \) where \( \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \Omega \) for \( s \leq t \).

(iii) A stochastic process \( \{ X(t) : t \geq 0 \} \) is a family of random variables all defined on the same probability space \( (\Omega, \mathcal{F}, \mathcal{P}) \), where \( \mathcal{F} \) is a history and \( \mathcal{P} \) is a probability measure on \( \Omega \).

(iv) A stochastic process \( X(t) \) is adapted to a history \( \mathcal{F}_t \) if \( X(t) \) is a function of \( \mathcal{F}_t \) for each \( t \), meaning at time \( t \) we know the outcome of \( X(s) \) for \( s \leq t \).
(v) A sample path is a sequence of outcomes (or realization) obtained from the stochastic process $X(t)$.

(vi) A predictable process $H(t)$ is process that is adapted to the history $\mathcal{F}_t$ and whose sample paths are left-continuous.

In non-mathematical terms, $\sigma$-algebra is a family of events that can be determined to have occurred by observing the past events of the experiment, and the history $\{\mathcal{F}_t : t \geq 0\}$ as defined above, which is all of the information concerning what has occurred for subjects in the experiment up to time $t$.

An important class of stochastic processes that serves as a foundation for many results in counting processes is called a martingale. For the purposes of this research we will focus on continuous time martingale processes. See Aalen et al. (2008) for more information on discrete time martingale concepts. Let $M = \{M(t) : t \geq 0\}$ be a process that is adapted to the history $\{\mathcal{F}_t\}$. The stochastic process $M$ is a martingale if the following properties are satisfied:

1. $\mathbb{E}[M(t) | \mathcal{F}_s] = M(s)$ for all $s < t$

2. $\mathbb{E}[dM(t) | \mathcal{F}_{t-}] = 0$

where $dM(t)$ denotes the increment of the process $M$ over a small interval $[t, t+dt]$. Essentially these properties state that the expected value of the martingale process $M$ in the future is its present value at time $s$, and the martingale process is a zero-drift process when adapted to some history. Throughout this dissertation
we will assume that $M(0) = 0$. Using double expectations, it can be shown that $E[M(t)] = 0$ for all $t$, meaning that $M$ is a mean zero martingale. It can also be shown that martingales have uncorrelated increments: $\text{Cov}(M(t) - M(s), M(v) - M(u)) = 0$ for all $0 < s < t < u < v$. Another type of martingale that will be utilized is the submartingale. Replacing the first property of martingales above, a process $M$ is a submartingale if it satisfies

$$E[M(t) | \mathcal{F}_s] \geq M(s) \text{ for all } t > s. \quad (2.31)$$

In essence this is saying that a submartingale is a nondecreasing process that increases as $t$ grows larger, indicating progression of time.

In addition to the properties using the expected value of a martingale, the variation of a martingale process is also important in applications. The predictable variation process $\langle M \rangle(t)$ is defined as the limit of the sum of conditional variances of the martingale differences:

$$\langle M \rangle(t) = \int_0^t \text{Var} [dM(u) | \mathcal{F}_{u-}] = \lim_{n \to \infty} \sum_{k=1}^n \text{Var} (\Delta M_k | \mathcal{F}_{(k-1)t/n}) \quad (2.32)$$

where the interval $[0, t]$ is partitioned into $n$ subintervals of length $t/n$ and $\Delta M_k = M(kt/n) - M((k - 1)t/n)$ is the increment of the martingale over the $k$th of these intervals. Note that the predictable variation process is a nondecreasing predictable process. It can be shown that the predictable variation process provides
an unbiased estimator of \( \text{Var}[M(t)] \) where \( M(t) \) is a mean-zero martingale. It can be shown that the increment \( d \langle M \rangle(t) \) of the predictable variation process over the small time interval \([t, t+dt]\) is the conditional variance of the increment \( dM(t) \) of the martingale \( M(t) \):

\[
d \langle M \rangle(t) = \text{Var}[dM(t)|\mathcal{F}_{t-}]
\]  

(2.33)

This result is key in deriving the variances for test statistics and other estimators that involve counting processes.

Up to this point we have addressed theoretical results in the case of a single martingale. However for the purposes of this research we need to account for multiple martingales. In general counting process theory, the \textit{predictable covariation process} of two martingales \( M_i(t) \) and \( M_j(t) \) where \( i \neq j \) is the following:

\[
\langle M_i, M_j \rangle(t) = \int_0^t \text{Cov}[M_i(t), M_j(t)|\mathcal{F}_{t-}]
\]

(2.34)

\[
d \langle M_i, M_j \rangle(t) = \text{Cov}[dM_i(t), dM_j(t)|\mathcal{F}_{t-}]
\]

(2.35)

In the case of \( n \) martingales, let \( M_i(t), i = 1, ..., n \) be martingales with respect to the same history \( \mathcal{F}_t \). Then the sum of the martingales \( M_\ast(t) = \sum_{i=1}^n M_i(t) \) is also a martingale with respect to the history \( \mathcal{F}_t \). Just like the rules for determining the variation of a linear combination of random variables, the predictable variation
process of a linear combination of martingales is found to be the following:

$$\langle M_t(t) \rangle = \sum \langle M_i \rangle (t) + 2 \sum \langle M_i, M_j \rangle (t)$$  (2.36)

When the predictable covariation process of two martingales $\langle M_i, M_j \rangle (t) = 0$ for all $t$, then $M_i(t)$ and $M_j(t)$ are called orthogonal martingales, and $\langle M_\cdot(t) \rangle = \sum \langle M_i \rangle (t)$.

Another important situation to consider for the development of estimators and test statistics is the transformation of a martingale process. Given a predictable process $H(t)$, the stochastic integral $I(t)$ is defined as the following:

$$I(t) = \int_0^t H(s) dM(s)$$  (2.37)

It can be shown that $I(t)$ is indeed a mean-zero martingale with respect to the history $\mathcal{F}_t$. Also, the predictable variation process of a stochastic integral is defined as the following:

$$\left\langle \int H dM \right\rangle = \int H^2 d\langle M \rangle$$  (2.38)

One of the most important theorems applicable to counting process theory is the Doob-Meyer decomposition. It states that any submartingale $Z(t)$ adapted
to the history $\mathcal{F}_t$ can be uniquely decomposed into a nondecreasing predictable process called a compensator $X(t)$ and a mean-zero martingale $M(t)$:

$$Z(t) = X(t) + M(t) \quad (2.40)$$
$$dZ(t) = dX(t) + dM(t) \quad (2.41)$$

The Doob-Meyer decomposition theorem ensures that for any submartingale there is a unique compensator (predictable process) that can be subtracted from the submartingale to produce a martingale. Two important consequences of the Doob-Meyer decomposition theorem are the following results:

1. $dX(t) = \mathbb{E}[dZ(t) | \mathcal{F}_t]$  
2. $dM(t) = dZ(t) - \mathbb{E}[dZ(t) | \mathcal{F}_t]$  

2.4.2 Counting Process Preliminary Concepts

Survival analysis and more specifically competing risks analysis are just two of the applications counting process theory can be applied to. We will first consider counting process concepts from a single subject perspective, and then move to multiple subjects. First we begin with the definition of a counting process and an intensity process.

Definition 2.4.2.

(i) A counting process $\{N(t) : t \geq 0\}$ is a stochastic process adapted to a history
\{F_t : t \geq 0\} with \(N(0) = 0\) and \(N(t) < \infty\) almost surely, and whose paths are with probability one right-continuous, piecewise constant, and have only jump discontinuities with jumps of size 1 at event times and are constant in between.

(ii) The *intensity* of a process is the conditional probability that an event occurs in the interval \([t, t+dt]\), given all that has been observed prior to this interval (the history of the process \(F_{t-}\)) divided by the length of that interval.

(iii) An *intensity process* \(\lambda(t)\) of a counting process specifies the probability of an event occurring in the time interval \([t, t + dt]\) given the history of the counting process \(F_{t-}\).

Let \(dN(t)\) denote the increment of the counting process \(N(t)\) in the time interval \([t, t + dt]\). The intensity process for the counting process \(N(t)\) is defined as

\[
\lambda(t) = \frac{P(dN(t) = 1|F_{t-})}{dt}
\]

\[
\lambda(t)dt = P(dN(t) = 1|F_{t-}) = E[dN(t)|F_{t-}]
\]

Since this is a nondecreasing process and adapted to the history \(F_t\), it is also a submartingale defined in 2.31. Applying the Doob-Meyer decomposition
theorem yields the following:

\[ N(t) = \Delta(t) + M(t) \]  \hspace{1cm} (2.42)

\[ dN(t) = d\Delta(t) + dM(t) \]  \hspace{1cm} (2.43)

where \( \Delta(t) \) is a nondecreasing predictable process, hence the compensator of the process \( N(t) \) and \( M(t) \) is a mean-zero martingale. In the special situation of counting processes, the increment of the predictable process \( d\Delta(t) \) is equal to the intensity process \( \lambda(t)dt \) due to the first result of the Doob-Meyer decomposition theorem:

\[ d \langle M \rangle (t) = \lambda(t)dt \]  \hspace{1cm} (2.44)

\[ \langle M \rangle (t) = \Delta(t) \]  \hspace{1cm} (2.45)

The predictable variation process for the counting process martingale \( M(t) \) is the following:

\[ \langle M(t) \rangle = \Delta(t) \]  \hspace{1cm} (2.46)

Development of estimators and test statistics often involve a transformation of the counting process martingale. Using the definition of a stochastic integral
and the Doob-Meyer decomposition, we have the following:

\[ I(t) = \int_0^t H(s) dN(s) - \int_0^t H(s) \lambda(s) ds \]  \hspace{1cm} (2.47)

While the second integral \( \int_0^t H(s) \lambda(s) ds \) turns out to be an ordinary integral, the first integral is the sum of the values of the predictable process \( H(t) \) at the jump times of the counting process \( N(s) \), a key result that is utilized in the derivation of the Nelson-Aalen estimator. Also, the predictable variation process of a stochastic integral of a counting process martingale is

\[
\left\langle \int_0^t HdM \right\rangle (t) = \int_0^t H^2(s) \langle M \rangle (s) = \int_0^t H^2(s) \lambda(s) ds
\]

In the case of \( n \) subjects, there are \( N_i(t), i = 1, \ldots, n \) counting processes associated with each subject with corresponding counting process martingales \( M_i(t), i = 1, \ldots, n \) and intensity processes \( \lambda_i(t), i = 1, \ldots, n \). It is assumed that no two of these counting processes can jump at the same time (we are dealing with continuous-time counting processes). A superposed counting process \( \tilde{N}(t) \) can be defined as the summation of the \( n \) individual counting processes:

\[
\tilde{N}(t) = \sum_{i=1}^{n} N_i(t) \]  \hspace{1cm} (2.48)
Since $N(t)$ is also a counting process, it follows that the associated intensity process $\lambda(t)$ is

$$\lambda(t) = \sum_{i=1}^{n} \lambda_i(t)$$  \hspace{1cm} (2.49)

Another important result for multiple counting processes is the summation of transformations. Let $M_i(t), i = 1, \ldots, n$ denote a collection of martingales with respect to the common history $\{\mathcal{F}_t, t \geq 0\}$ and $H_i(t), i = 1, \ldots, n$ denote a collection of predictable processes with respect to the common history $\{\mathcal{F}_t, t \geq 0\}$. Then $\sum_i \int H_i dM_i$ is a martingale with respect to the history $\{\mathcal{F}_t, t \geq 0\}$. Many of the estimators and test statistics used in event history analysis with censored data are written as forms of this summation where $M_i(t), i = 1, \ldots, n$ are martingales formed by observed counting process data $dN(t)$ and compensator $\lambda(t)$ (Therneau & Grambsch, 2001).

Due to the construction of the counting process martingale, it is clear that no two of the counting process martingales can jump at the same time either. Hence, for all $i \neq j$,

$$\langle M_i, M_j \rangle (t) = \text{Cov}[M_i(t), M_j(t) | \mathcal{F}_{t-}] = 0 \text{ for all } t.$$  \hspace{1cm} (2.50)

It is clear that the martingales are orthogonal, thus we have the sum of predictable
variation processes to be the following:

\[
\left\langle \sum_{i=1}^{n} \int H_j dM_j \right\rangle (t) = \sum_{i=1}^{n} \int_{0}^{t} H^2(s) \lambda(s) ds
\]

2.4.3 Application to Survival Analysis

This section applies the counting process and martingale theory from the previous sections to estimation of cumulative hazards in survival data. Assume we have \( N_i(t) = I \{T_i \leq t\}, i = 1, ..., n \) counting processes associated with the \( i \)th subject. Along with the observed counting process, we also observe an at-risk process for each subject denoted by \( Y_i(t) = I \{T_i > t\} \) that takes into account possible censoring. Both of these processes are adapted to the history \( \mathcal{F}_t \) that contains the history of all of the counting processes as well as the risk processes and any covariates. The intensity process \( \lambda_i(t) \) is the following:

\[
\lambda_i(t) = E[dN_i(t)|\mathcal{F}_{t-}]
\]

\[
= Y_i(t)h_i(t)dt
\]

where \( h_i(t) \) is the hazard function for the \( i \)th process. Assuming that the survival times for the subjects are independently identically distributed with hazard
function $h(t)$, then we have the multiplicative intensity model given by

$$\lambda_i(t) = Y_i(t)h(t)dt$$  \hfill (2.51)

Using superposed counting processes and risk processes, the superposed intensity function $\lambda(t)$ is the following:

$$\lambda(t) = E[d\tilde{N}(t)|\mathcal{F}_{t-}]$$

$$= \tilde{Y}(t)h(t)dt$$

Applying the Doob-Meyer decomposition to the process $\tilde{N}(t)$ results in the following representation:

$$d\tilde{N}(t) = \tilde{Y}(t)h(t)dt + dM(t)$$  \hfill (2.52)

Dividing both sides by $\tilde{Y}(t)$ and introducing the process $J(t) = I(\tilde{Y}(t) > 0)$ yields the following:

$$\frac{J(t)}{\tilde{Y}(t)} = J(t)h(t)dt + \frac{J(t)}{\tilde{Y}(t)}dM(t)$$  \hfill (2.53)

In general, it is difficult to derive an estimator for $h(t)dt$ due to being a nonparametric quantity. However, integrating both sides provides a way to estimate the
cumulative hazard:

$$\int_0^t \frac{J(s)}{Y(s)} d\bar{N}(s) = \int_0^t J(s)h(s)ds + \int_0^t \frac{J(s)}{Y(s)} dM(s) \tag{2.54}$$

Due to the fact that the counting process $\bar{N}(t)$ is a step function with jumps of size 1 at event times $T_1 < T_2 < \ldots < T_n$, the integral on the left side reduces to the following:

$$\int_0^t \frac{J(s)}{Y(s)} d\bar{N}(t) = \sum_{T_i \leq t} \frac{1}{Y(T_i)} \tag{2.55}$$

which is the Nelson-Aalen estimate (denoted as $\hat{A}(t)$) of the cumulative hazard. Maximum likelihood estimation techniques can also be used to derive the Nelson-Aalen estimate, but using the counting process and martingale theory allows for unified techniques to derive variances, confidence intervals, and testing procedures for this and other kinds of estimators.

The variance of the Nelson-Aalen estimator is the following:

$$\sigma^2(t) = \int_0^t \frac{J(s)}{Y^2(s)} d\bar{N}(t) \tag{2.56}$$

To justify this quantity, let $\bar{A}(t) = \int_0^t \frac{J(s)}{Y(s)} d\bar{N}(s)$ and $A^*(t) = \int_0^t J(s)h(s)ds$. Computing the difference between these two processes yields a stochastic integral as
well as mean-zero martingale:

$$A(t) - A^*(t) = \int_0^t \frac{J(s)}{Y(s)} dM(s)$$

(2.57)

Hence, using the rule for the predictable variation process of a stochastic integral, the estimated variance of the Nelson-Aalen estimate is the following:

$$\hat{\sigma}^2(t) = \int_0^t \frac{J(s)}{Y^2(s)} d\tilde{N}(s)$$

(2.58)

Similar to the Nelson-Aalen estimator, the integral on the left side reduces to

$$\hat{\sigma}^2(t) = \sum_{T_i \leq t} \frac{1}{Y^2(T_i)}$$

(2.59)

2.4.4 Application to Competing Risks Analysis

The results of applying counting process and martingale theory to survival data is easily extended to competing risks data. Let $N_i(t) = \{N_{i1}(t), ..., N_{ip}(t)\}$, $i = 1, ..., n$ denote the multivariate counting process for the $i$th subject, where $N_{ij}(t) = I(T_i \leq t, \delta_i = j)$ corresponds to the counting process of the $j$th competing risk for the $i$th subject. The risk process for the $i$th subject is $Y_i(t) = I(T_i \geq t)$, meaning the subject is at risk until either one of the competing risk events or censoring occurs at time $T_i$. Both of these processes are adapted to the history $\mathcal{F}_{it}$ that contains the history of the $p$ competing risks counting processes for subject
as well as their risk processes and any covariates. The intensity process \( \lambda_{ij}(t) \) is the following:

\[
\lambda_{ij}(t) = \mathbb{E}[dN_{ij}(t) | \mathcal{F}_t]
\]

\[
= Y_i(t) h_j(t) dt
\]

where \( h_j(t) \) is the cause-specific hazard function for the \( j \)th competing risk.

Before addressing the case of multiple subjects, here are some definitions of superposed processes used in the derivations:

1. \( N_j(t) = \sum_{i=1}^{n} N_{ij}(t) \)

2. \( \bar{Y}(t) = \sum_{i=1}^{n} Y_i(t) \)

3. \( \bar{N}(t) = \sum_{i=1}^{n} N_j(t) \)

4. \( M_{ij}(t) = N_{ij}(t) - \int_0^t Y_i(s) h_j(s) ds \)

5. \( M_j(t) = \sum_{i=1}^{n} M_{ij}(t) \)

6. \( \mathcal{F}_t = \sigma\{\bigcup_{i=1}^{n} \mathcal{F}_{it}\} \)

Note that \( M_{ij}(t) \) is a martingale with respect to the history \( \mathcal{F}_{it} \). Using the fact that the sum of martingales is also a martingale in the counting process situation, we can see that

\[
M_j(t) = N_j(t) - \int_0^t \bar{Y}(s) h_j(s) ds \quad (2.60)
\]
is a martingale with respect to the history $\mathcal{F}_t$ for the $j$th competing risk. Using the same structure in the proof of the Nelson-Aalen estimator in the survival data case, the Nelson-Aalen estimator $\hat{A}_j(t)$ of the cumulative cause-specific hazard for the $j$ competing risk is the following:

$$\int_0^t \frac{J(s)}{\bar{Y}(s)} dN_j(t) = \sum_{T_{ij} \leq t} \frac{1}{\bar{Y}(T_{ij})}$$

(2.61)

where $T_{1j} < T_{2j} < \ldots < T_{nj}$ are the failure times for the $j$th competing risk event. Similarly, the variance estimate for the Nelson-Aalen estimator for the $j$ competing risk is the following:

$$Var[\hat{A}_j(t)] = \sum_{T_{ij} \leq t} \frac{1}{\bar{Y}^2(T_{ij})}$$

(2.62)

It follows that an estimate of the cumulative incidence function for risk $j$ is the following:

$$\hat{F}_j(t) = \sum_{T_k \leq t} \hat{S}(T_{k-1}) \hat{A}_j(T_k)$$

(2.63)

where $\hat{S}(T_{k-1})$ is the Kaplan-Meier estimate for the probability of remaining event free prior to time $T_k$. 
CHAPTER III

INFERENCE PROCEDURES

3.1 Introduction

The primary goal of this research is to determine whether the reasons of discontinuation are equally serious or if some reasons are more serious than others, in the clinical trial setting in which the risks of discontinuation due to different reasons are acting simultaneously. This chapter details the two inference procedures used in this dissertation to detect differences between competing risks. Section 3.2 discusses the powerful martingale central limit theorem that is used to justify the approximate null distributions of test statistics. Section 3.3 details the KLY test procedure used to test two competing risks. Section 3.4 discusses the TKC test procedures and its three subtests which test whether competing risks are significantly different within different groups, whether groups are significantly different within competing risks, and whether competing risk and group combinations are significantly different.

3.2 Martingale Central Limit Theorem

In many statistical models, the central limit theorem plays a key role in the derivation of asymptotic distributions for test statistics. In counting process and
martingale theory, there is also a version of the central limit theorem that plays the same role. There are several versions of the martingale central limit theorem, and we use the version proposed by Rebolledo (1980) and discussed in Anderson et al (1993) and Fleming & Harrington (1991). This powerful theorem is built on the transformations of a stochastic process called the Brownian motion:

**Definition 3.2.1.** The standard Brownian motion process $W(t)$ is a stochastic process with the following properties:

- $W(t)$ has continuous sample paths, i.e. the values taken by $W$ over the time period $[0, t]$ are continuous;
- $W(0) = 0$ and $\mathbb{E}[W(0)] = 0$ for any $t$;
- $W(t)$ has independent increments, meaning for any $0 \leq u \leq t$, $W(t) - W(u)$ is independent of $W(u)$;
- $W(t)$ has variance equal to $t$;
- $W(t)$ is a Gaussian process, meaning for any positive integer $n$ and sequence $t_1, ..., t_n$, the joint distribution of $\{W(t_1), W(t_2), ..., W(t_n)\}$ is distributed as a multivariate normal.

Let $f$ be a measurable nonnegative function and let $\alpha(t) = \int_0^t f^2(s)ds$. Then $\int f dW$ denotes a process that satisfies property 2, 3, and 5, but with

$$\text{Var} \left[ \int_0^t f(s)dW(s) \right] = \alpha(t) \quad (3.1)$$
The following is a multivariate version of the martingale central limit theorem by Fleming & Harrington (1991):

**Theorem 3.2.1** (Multivariate Martingale Central Limit Theorem). Let $W_1, \ldots, W_r$ be $r$ independent Brownian motion processes and $f_1, \ldots, f_r$ be $r$ measurable non-negative functions such that $\alpha_l(t) = \int_0^t f_l^2(s) ds < \infty$ for all $t > 0$ and $l = 1, \ldots, r$.

Assume that for each $l = 1, \ldots, r$ that, as $n \to \infty$,

$$
(U_l) \overset{P}{\to} \int_0^t f_l^2(s) ds
$$

Then

$$
(U_1, \ldots, U_r) \Rightarrow \left( \int f_1 dW_1, \ldots, \int f_r dW_r \right)
$$

Modifications of the Brownian motion provide models for underlying processes, and in particular time-transformed Brownian motions serve as limits of stochastic integrals of counting process martingales. Let $W^*(t) = W(V(t))$ denote the time-transformed Brownian motion, and $V(t)$ denote a strictly increasing continuous function where $V(0) = 0$. Then $W^*(t)$ is also a stochastic process inheriting the properties of $W(t)$, with

- $W^*(t)$ has continuous sample paths;
- $W^*(0) = 0$ and $E[W^*(0)] = 0$ for any $t$;
- $W^*(t)$ has independent increments;
• $W^*(t)$ has variance equal to $V(t)$;

• $W^*(t)$ is a Gaussian process.

• $W^*(t)$ is a mean-zero martingale with predictable variation process $<W^*(t)> = V(t)$

In counting process theory, many of the test statistics are sums of stochastic integrals taking the form

$$Z(t) = \sum_{j=1}^{k} \int_{0}^{t} H_{j}^{(n)}(s) dM_{j}^{(n)}(s)$$  \quad (3.4)$$

where $H_{j}^{(n)}$ is a predictable process for each $n$ and

$$M_{j}^{(n)}(t) = N_{j}^{(n)}(t) - \int_{0}^{t} \lambda_{j}^{(n)}(s) ds$$  \quad (3.5)$$

is a counting process martingale for $j = 1, ..., k$. Assume that $V(t) = \int_{0}^{t} v(s) ds$ is a strictly increasing continuous function where $V(0) = 0$. There are two conditions (apart from some regulatory conditions which are not the focus here) that must be satisfied in order for a summation of stochastic integrals to converge in distribution to a time-transformed Brownian motion:

1. $\sum_{j=1}^{k} (H_{j}^{(n)}(s))^2 \lambda_{j}^{(n)}(s) ds \overset{p}{\to} v(s) > 0$ for all $s \in [0, \tau]$ as $n \to \infty$

2. $H_{j}^{(n)}(s) \overset{p}{\to} 0$ for all $j = 1, ..., k$ and $s \in [0, \tau]$ as $n \to \infty$
Condition one ensures that the predictable variation processes of the martingales converge to some deterministic function, and condition two ensures that the size of the jumps of the martingales converge to zero. When these two conditions are met, then the sum of stochastic integrals $Z(t) \overset{p}{\rightarrow} W(V(t))$, where $W(V(t))$ has mean zero and variance $V(t)$. Using standardization, $Z(t)/V(t) \overset{p}{\rightarrow} W(t)$.

### 3.3 KLY Test Procedure Comparing Competing Risks

The KLY test procedure was developed by Kochar, Lam, and Yip (2002). Previous versions of test statistics along with theoretical results used here are found in Aly, Kochar & McKeague (1994). In the situation of two competing risk events, we wish to test the following hypotheses:

$$H_0 : h_1(t) = h_2(t), \text{ for all } t \quad (3.6)$$

against the alternative

$$H_1 : h_1(t) \neq h_2(t), \text{ for all } t, \text{ with strict inequality for some } t. \quad (3.7)$$

These hypotheses can be stated using cumulative incidence functions instead:

$$H_0 : F_1(t) = F_2(t), t \geq 0 \quad (3.8)$$
against the alternative

\[ H_1 : F_1(t) \neq F_2(t), \forall t \geq 0, \text{ with strict inequality for some } t \quad (3.9) \]

due to the fact that \( F_j(t) = \int_0^t S(u)h_j(u)du \). The test statistic used is a form of a weighted log-rank type statistic:

\[
L_n(t) = \int_0^t w(s) \left[ d\hat{H}_2(s) - d\hat{H}_1(s) \right] \tag{3.10}
\]

where \( w(s) \) is a weight function that reflects the importance attached to the difference between the cause-specific hazard functions at time \( s \) and \( d\hat{H}_j(r) = dN_j(t)/\bar{Y}(t) \). In fact, \( L_n(t) \) is a mean-zero martingale under \( H_0 \).

Proof.

\[
L_n(t) = \int_0^t w(u) \left[ d\hat{H}_2(u) - d\hat{H}_1(u) \right] \\
= \int_0^t w(u) \left[ \frac{dN_2(u)}{\bar{Y}(u)} - \left( \frac{dN_1(u)}{\bar{Y}(u)} \right) \right] \\
= \int_0^t w(u) \left[ \frac{dM_2(u) + h(u)Y(u)}{\bar{Y}(u)} - \left( \frac{dM_1(u) + h(u)Y(u)}{\bar{Y}(u)} \right) \right] \\
= \int_0^t \frac{w(u)}{\bar{Y}(u)} \left[ dM_2(u) + h(u)Y(u) - dM_1(u) - h(u)Y(u) \right] \\
= \int_0^t \frac{w(u)}{\bar{Y}(u)} [dM_2(u) - dM_1(u)].
\]
Note that $M_1(t)$ and $M_2(t)$ are mean-zero martingales, so it follows that $L_n(t)$ is a mean-zero martingale under $H_0$. □

To help meet the conditions required by the martingale central limit theorem, we concentrate on $\sqrt{n}L_n(t)$ which has a predictable variation process

$$
\sigma_n^2(t) = \int_0^t \frac{nw^2(u)}{Y^2(u)} (dH_2(u) + dH_1(u))
$$

Also, the variance of $\sqrt{n}L_n(t)$ is estimated by the following:

$$
S_n^2(t) = \int_0^t \frac{nw^2(s)}{Y^2(s)} \, d\bar{N}(s), \text{ where } d\bar{N}(t) = dN_1(t) + dN_2(t).
$$

Proof. Recall that $L_n(t) = \int_0^t w(s) \left[ d\hat{H}_2(s) - d\hat{H}_1(s) \right]$. $L_n(t)$ is a stochastic integral because $w(t)$ is a predictable process and the quantity $d\hat{H}_2(s) - d\hat{H}_1(s)$ is a martingale. Using the definition of a predictable variation process of a stochastic integral:

$$
\langle L_n(t) \rangle = \int_0^t w^2(s) d\left( \hat{H}_2(s) - \hat{H}_1(s) \right)
$$

Using the fact that $M_1(t)$ and $M_2(t)$ are orthogonal martingales, we have the
The following theorem from Aly, Kochar, and McKeague (1994) addresses the convergence of the statistic $\sqrt{n}L_n(t)$:

**Theorem 3.3.1.** Let $w$ be a locally bounded predictable non-negative weight function such that $nw^2(t)/\bar{Y}(t) \rightarrow K(t)$ in probability for each $t$ and $\int_0^\infty K(u)(d\Delta_2 + d\Delta_1) < \infty$. Then under $H_0$, $\sqrt{n}L_n(t) \overset{D}{\rightarrow} W(\sigma^2(t))$ where $W(t)$, $t \geq 0$ is a standard Brownian motion and $\sigma^2(t) = \int_0^\infty K(u)(dH_2 + dH_1)$, which can be estimated consistently by $S_n^2(t)$.

**Proof.** Under the null hypothesis, we have the following:

$$\sqrt{n}L_n(t) = \sum_{j=1}^{k=2} \int_0^t H_j^{(n)}(s) dM_j(s) \quad (3.14)$$
where

\[ H_1^{(n)}(t) = \frac{w(t)}{Y(t)\sqrt{n}} \quad \text{and} \quad H_2^{(n)}(t) = -\frac{w(t)}{Y(t)\sqrt{n}} \]  

(3.15)

are predictable processes for each \( n \). Recall that the martingale \( M_j(t) \) can be written as the following:

\[
M_j(t) = N_j(t) - \hat{H}_j(t)
\]

\[
= N_j(t) - \int_0^t \lambda_j(t)
\]

\[
= N_j(t) - \int_0^t \bar{Y}(s)h_j(s)ds
\]

Recall that two conditions need to be met in order for a test statistic like \( L_n(t) \) to converge to a time-transformed Brownian motion:

1. \( \sum_{j=1}^k (H_j^{(n)}(s))^2 \lambda_j^{(n)}(s)ds \to v(s) > 0 \) for all \( s \in [0, \tau] \) as \( n \to \infty \)

2. \( H_j^{(n)}(s) \to 0 \) for all \( j = 1, \ldots, k \) and \( s \in [0, \tau] \) as \( n \to \infty \)

Applying to \( L_n(t) \), the first condition can be explored with the following:

\[
\sum_{j=1}^{k=2} H_j^{(n)}(t)\lambda_j(t) = \left( \frac{\sqrt{n}w(t)}{\bar{Y}(t)} \right)^2 \bar{Y}(t)h_2(t) + \left( -\frac{\sqrt{n}w(t)}{\bar{Y}(t)} \right)^2 \bar{Y}(t)h_1(t)
\]

\[
= \frac{nw^2(t)}{\bar{Y}(t)} h_2(t) + \frac{nw^2(t)}{\bar{Y}(t)} h_1(t)
\]

\[
= \frac{nw^2(t)}{\bar{Y}(t)} (h_2(t) + h_1(t))
\]
Letting \( \frac{n w^2(t)}{Y(t)} \) \( \overset{p}{\rightarrow} \) \( K(t) \) for each \( t \), it follows that

\[
\frac{n w^2(t)}{Y(t)} (h_2(t) + h_1(t)) \overset{p}{\rightarrow} K(t)(h_2(t) + h_1(t)) \tag{3.16}
\]

and condition one is satisfied. Note that the second condition is also satisfied due to the following:

\[
\sqrt{n} \frac{w(t)/n}{Y(t)/n} \overset{p}{\rightarrow} 0 \tag{3.17}
\]

The test statistic used in this research is the following:

\[
C^*_n = \sup_{0 \leq s < t} \frac{|L_n(t) - L_n(s)|}{S_n(\infty)} \text{ for } H_1 : F_1(t) \neq F_2(t) \tag{3.18}
\]

Due to the continuous mapping theorem,

\[
C^*_n \overset{p}{\rightarrow} \sup_{0 \leq s < t} \frac{|W(t) - W(s)|}{S_n(\infty)} = C^{**} \tag{3.19}
\]

where \( W(t) \) is the standard Brownian motion. Then \( C^{**} \) has the same distribution as the range of the standard Brownian motion \( (||W^-|| + ||W^+||) \), where \( ||W^-|| = \min(0, \inf W(t)) \) and \( ||W^+|| = \max(0, \sup W(t)) \). Feller (Feller, 1951) derived the
probability density function of the range as the following:

$$f(x) = 8 \sum_{k=1}^{\infty} (-1)^{k-1} k^2 \phi(kx). \quad (3.20)$$

In simulation studies conducted by the authors, tests based on $C_n^*$ were robust for a wide selection of weight functions. The weight function $w(t) = \bar{Y}(t)$ (the total number of subjects at risk at time $t$) will be used in the application of the KLY test in chapter V.

3.4 TKC Tests

3.4.1 Introduction

The TKC test, established by Tiwari et al (Tiwari et al., 2006) extends the KLY test by allowing for more than two competing risks to be compared and for comparisons among different groups. The authors developed a test that simultaneously determines whether there are differences between the CSH between competing risks and groups, which achieves the desired nominal level of significance as opposed to conducting multiple tests to compare competing risks and groups. The authors also show that the TKC test as well as separate tests that compare competing risks such as the test proposed by Lam (Lam, 1998), and tests that compare groups within competing risks such as the test proposed by Kulathinal and Gasbarra (Kulathinal & Gasbarra, 2002) can be formulated using
a unified application of counting process and martingale theory. Let $j = 1, \ldots, J$ denote the event type and let $k = 1, \ldots, K$ denote the group. Let $H_{01}$ denote the null hypothesis for the first TKC subtest comparing all competing risk of discontinuation and group level combinations, $H_{02}$ denote the null hypothesis for the second TKC subtest comparing groups within each competing risk of discontinuation, and let $H_{03}$ denote the null hypothesis for the third TKC subtest comparing competing risks of discontinuation within each group. The null hypotheses for these subtests of the TKC procedure are formulated as the following:

$$H_{01} : h_{jk}(t) = h_{00}(t) \text{ vs } H_{11} : h_{jk}(t) \neq h_{00}(t) \text{ for all } j \text{ and for all } t,$$

$$H_{02} : h_{jk}(t) = h_{j0}(t) \text{ vs } H_{12} : h_{jk}(t) \neq h_{j0}(t), \quad k = 1, \ldots, K \text{ for all } j \text{ and for all } t,$$

$$H_{03} : h_{jk}(t) = h_{0k}(t) \text{ vs } H_{13} : h_{jk}(t) \neq h_{0k}(t), \quad j = 1, \ldots, J \text{ for all } k \text{ and for all } t.$$

Let $T_{jk}$ denote the random variable associated with the time to discontinuation due to reason $j$ and group $k$. Let $n_k$ denote the number of subjects in group $k$, $k = 1, \ldots, K$. Let $n = \sum_{k=1}^{K} n_k$ denote the number of subjects with censored and uncensored event times $t_1, t_2, \ldots, t_n$. Let $0 < t_{(1)} < t_{(2)} < \cdots < t_{(r)}$ be the $r$ ordered distinct event times. The counting process for the number of subjects who discontinue due to reason $j$ in group $k$ and the risk process for the number
of subjects still at risk for discontinuation prior to time $t$ in group $k$ be defined as

$$N_{jk}(t) = \sum_{i=1}^{n_k} I[T_{ik} \leq t, \delta_{ik} = j], j = 1, \ldots, J, k = 1, \ldots, K$$

$$Y_k(t) = \sum_{i=1}^{n_k} I[T_{ik} \geq t], k = 1, \ldots, K$$

Similar to the definition of the CCSH in 2.19, the CCSH for discontinuation type $j$ and group $k$ and its associated Nelson-Aalen estimate are given by

$$H_{jk}(t) = \int_0^t h_{jk}(s) ds, \quad (3.21)$$

$$\hat{H}_{jk}(t) = \int_0^t \frac{dN_{jk}(s)}{Y_k(s)} \quad (3.22)$$

Recall that the counting process $N_{jk}(t)$ is a step function and the Nelson-Aalen estimate of $H_{jk}(t)$ can be represented using summation notation as

$$\hat{H}_{jk}(t) = \sum_{T_{jk} \leq t} \frac{1}{Y_k(T_{jk})} . \quad (3.23)$$

For the rest of this chapter, the results will be presented in the counting process integral notation.

The subtests are conducted by comparing estimated CCSH's for each competing risk and group combination to their estimated standard levels under the null hypothesis of the particular test. Under $H_{01}, H_{02},$ and $H_{03},$ the Nelson-Aalen
estimates of the CCSH’s are estimated by

\[
\begin{align*}
\hat{H}_{00}(t) &= \int_0^t \frac{dN_{**}(s)}{JY_*(s)}, \\
\hat{H}_{j0}(t) &= \int_0^t \frac{dN_{j*}(s)}{Y_*(s)}, \\
\hat{H}_{0k}(t) &= \int_0^t \frac{dN_{*k}(s)}{JY_k(s)},
\end{align*}
\]

where \(N_{**}(t) = \sum_{j=1}^J \sum_{k=1}^K N_{jk}(t)\), \(N_{j*}(t) = \sum_{k=1}^K N_{jk}(t)\), and \(N_{*k}(t) = \sum_{j=1}^J N_{jk}(t)\) are the aggregated counting processes used in the CCSH estimates. In the derivation of the null distributions of the test statistics, the associated martingales for these counting processes will be used via the martingale central limit theorem.

These martingale processes are the following:

\[
\begin{align*}
M_{**}(t) &= N_{**}(t) - \int_0^t JY_*(s)dH_{00}(s), \\
M_{j*}(t) &= N_{j*}(t) - \int_0^t Y_*(s)dH_{j0}(s), \\
M_{*k}(t) &= N_{*k}(t) - \int_0^t JY_k(s)dH_{0k}(s), \\
M_{jk}(t) &= N_{jk}(t) - \int_0^t Y_k(s)dH_{jk}(s)
\end{align*}
\]

Each of the hypothesis tests have a test process that measures the deviation between the CCSH estimates \(\hat{H}_{jk}(t)\) and the estimates of the common values of the CCSH’s under the corresponding null hypothesis. The test processes that compare the estimated CCSH for discontinuation type \(j\) and group \(k\) with the
estimated standard level under $H_{01}, H_{02}, H_{03}$ are the following:

$$Z_{jk}^{(1)} = \int_0^t W_{jk}(s) d\left[\hat{H}_{jk}(s) - \hat{H}_{00}(s)\right],$$

$$Z_{jk}^{(2)} = \int_0^t W_{jk}(s) d\left[\hat{H}_{jk}(s) - \hat{H}_{00}(s)\right],$$

$$Z_{jk}^{(3)} = \int_0^t W_{jk}(s) d\left[\hat{H}_{jk}(s) - \hat{H}_{0k}(s)\right],$$

where $W_{jk}$ is a predictable weight function. Like the KLY test, the weight function is used to reflect the importance attached to the difference between the estimated CCSH and the corresponding estimate under $H_0$. Weight functions are discussed in section 3.4.3.

### 3.4.2 Asymptotic Distributions

Under the assumption that the null hypotheses are true, the three test processes become the following:

$$\frac{1}{\sqrt{n}} Z_{jk}^{(1)} = \int_0^t \frac{W_{jk}(s)}{\sqrt{nY_k(s)}} dM_{jk}(s) - \int_0^t \frac{W_{jk}(s)}{\sqrt{nY_\bullet(s)}} dM_{\bullet\bullet}(s),$$

$$\frac{1}{\sqrt{n}} Z_{jk}^{(2)} = \int_0^t \frac{W_{jk}(s)}{\sqrt{nY_k(s)}} dM_{jk}(s) - \int_0^t \frac{W_{jk}(s)}{\sqrt{nY_\bullet(s)}} dM_{\bullet}(s),$$

$$\frac{1}{\sqrt{n}} Z_{jk}^{(3)} = \int_0^t \frac{W_{jk}(s)}{\sqrt{nY_k(s)}} dM_{jk}(s) - \int_0^t \frac{W_{jk}(s)}{\sqrt{nY_k(s)}} dM_{k}(s).$$

Each of these test processes now becomes a difference between two stochastic integrals and also a mean-zero martingale. The predictable variation and covariation
processes of the test processes are given by the following:

\[
\begin{align*}
\left\langle Z_{jk}^{(1)}, Z_{jk'}^{(1)} \right\rangle &= \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\delta_{jj'}\delta_{kk'}}{Y_k(s)} - \frac{1}{Y_\bullet(s)} \right) dH_{00}(s), \\
\left\langle Z_{jk}^{(2)}, Z_{jk'}^{(2)} \right\rangle &= \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \delta_{kk'} \frac{1}{Y_k(s)} - \frac{1}{Y_\bullet(s)} \right) dH_{j0}(s), \\
\left\langle Z_{jk}^{(3)}, Z_{j'k'}^{(3)} \right\rangle &= \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\delta_{jj'}}{Y_k(s)} - \frac{1}{Y_\bullet(s)} \right) \frac{1}{Y_k(s)} dH_{0k}(s)
\end{align*}
\]

where \( \delta_{uv} = 1 \) if \( u = v \) and \( \delta_{uv} = 0 \) otherwise. The derivation of \( \left\langle Z_{jk}^{(2)}, Z_{jk'}^{(2)} \right\rangle \) is detailed in the following proof:

**Proof.**

\[
Z_{jk}^{(2)} = \int_0^t W_{jk}(s) \left[ \hat{H}_{jk}(s) - \hat{H}_{j0}(s) \right]
\]

\[
= \int_0^t W_{jk}(s) \left[ \frac{dN_{jk}(s)}{Y_k(s)} - \frac{dN_{j\bullet}(s)}{Y_\bullet(s)} \right] - \int_0^t W_{jk}(s) \left[ \frac{dM_{jk}(s)}{Y_k(s)} + \frac{\sum_{k'=1}^K h_{j0}(s) Y_{k'}(s)}{Y_\bullet(s)} \right]
\]

\[
= \int_0^t W_{jk}(s) \left[ \frac{dM_{jk}(s)}{Y_k(s)} - \frac{dM_{j\bullet}(s)}{Y_\bullet(s)} \right] + \int_0^t W_{jk}(s) \left[ \frac{h_{j0}(s) Y_k(s)}{Y_k(s)} - \frac{\sum_{k'=1}^K h_{j0}(s) Y_{k'}(s)}{Y_\bullet(s)} \right]
\]

\[
= \int_0^t W_{jk}(s) \left[ \frac{dM_{jk}(s)}{Y_k(s)} - \frac{dM_{j\bullet}(s)}{Y_\bullet(s)} \right] + \int_0^t W_{jk}(s) \left[ h_{j0}(s) - \frac{h_{j0}(s) Y_\bullet(s)}{Y_\bullet(s)} \right]
\]

Using the result for the predictable variation process of a stochastic integral given in 2.38, the derivation of the predictable variation process for \( \left\langle \frac{M_{jk}(t)}{Y_k(t)} - \frac{M_{j\bullet}(t)}{Y_\bullet(t)} \right\rangle \) is
the following:

\[
\left\langle \frac{M_{jk}(t)}{Y_k(t)} - \frac{M_{j*}(t)}{Y_*(t)} \right\rangle = \left\langle M_{jk}(t) \left( \frac{1}{Y_k(t)} - \frac{1}{Y_*(t)} \right) - \frac{1}{Y_*(t)} \sum_{h \neq k} M_{jh}(t) \right\rangle \\
= \text{Var} \left[ M_{jk}(t) \left( \frac{1}{Y_k(t)} - \frac{1}{Y_*(t)} \right) \right] + \text{Var} \left[ \frac{1}{Y_*(t)} \sum_{h \neq k} M_{jh}(t) \right] \\
= \int_0^t \left( \frac{1}{Y_k(s)} - \frac{1}{Y_*(s)} \right)^2 h_{j0}(s) Y_k(s) ds + \int_0^t \frac{1}{(Y_*(s))^2} \sum_{h \neq k} h_{j0}(s) Y_h(s) ds \\
= \int_0^t \left( \frac{1}{Y_k(s)} - \frac{2}{Y_*(s)} + \frac{Y_k(s)}{(Y_*(s))^2} \right) h_{j0}(s) + \frac{1}{(Y_*(s))^2} \sum_{h \neq k} h_{j0}(s) Y_k(s) ds \\
= \int_0^t \left( \frac{1}{Y_k(s)} - \frac{2}{Y_*(s)} + \frac{1}{(Y_*(s))^2} \sum_{k=1}^K Y_k(s) \right) h_{j0}(s) ds \\
= \int_0^t \left( \frac{1}{Y_k(s)} - \frac{1}{Y_*(s)} \right) h_{j0}(s) ds
\]

Similarly, the predictable covariation process for \( k \neq k' \) is the following

\[
\left\langle \frac{M_{jk}(t)}{Y_k(t)} - \frac{M_{j*}(t)}{Y_*(t)}, \frac{M_{jk'}(t)}{Y_k'(t)} - \frac{M_{j*}(t)}{Y_*(t)} \right\rangle = -\int_0^t \frac{1}{Y_*(s)} h_{j0}(s) ds \quad (3.24)
\]

Combining the two results, the predictable variation and covariation processes for

\[\langle Z_{jk}^{(2)}, Z_{jk'}^{(2)} \rangle\]

is the following:

\[
\langle Z_{jk}^{(2)}, Z_{jk'}^{(2)} \rangle = \int_0^t W_{jk} W_{j'k'} \left( \frac{\delta_{kk'}}{Y_k(s)} - \frac{1}{Y_*(s)} \right) h_{j0}(s) ds \quad (3.25)
\]

To justify the convergence in distribution of the test processes, first are
some useful results due to the Glivenko-Cantelli theorem (Tiwari et al., 2006).

Let \( p_k = \lim_{n \to \infty} n_k/n, y_k = p_k S_k(t)S_U(t), \) and \( y_\bullet = \sum_k y_k. \)

- \( \| \frac{y_k}{n} - y_k \|_t \to 0 \)
- \( \| \frac{y_\bullet}{n} - y_\bullet \|_t \to 0 \)

where \( \| \cdot \|_a \) is the supremum norm over \([0, a]\). Using the martingale central limit theorem given in section 3.2, each of the test processes \( n^{-1/2}Z_{jk}^{(1)}, n^{-1/2}Z_{jk}^{(2)}, n^{-1/2}Z_{jk}^{(3)} \) converge to functions of Brownian motions \( B_{jk}^{(1)}, B_{jk}^{(2)}, B_{jk}^{(3)} \) respectively with the following asymptotic variance-covariance functions:

\[
\text{Cov} \left[ U_{jk}^{(1)}(t), U_{j'k'}^{(1)}(t) \right] = \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\delta_{jj'}\delta_{kk'}}{y_k} - \frac{1}{y_\bullet} \right) dH_{00}(s),
\]

\[
\text{Cov} \left[ V_{jk}^{(2)}(t), V_{j'k'}^{(2)}(t) \right] = \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\delta_{kk'}}{y_k} - \frac{1}{y_\bullet} \right) dH_{j0}(s),
\]

\[
\text{Cov} \left[ Q_{jk}^{(3)}(t), Q_{j'k'}^{(3)}(t) \right] = \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\delta_{jj'}}{y_k} - \frac{1}{y_\bullet} \right) \frac{1}{y_k} dH_{0k}(s)
\]

The derivation of \( \text{Cov} \left[ V_{jk}^{(2)}(t), V_{j'k'}^{(2)}(t) \right] \) is detailed in the following proof (Tiwari et al., 2006):

**Proof.** Assume that \( \| W_{jk}/n - W_{jk}^0 \|_t \to P 0 \) and recall that under \( H_0, n^{-1/2}Z_{jk}^{(2)} \) can be written as the following:

\[
n^{-1/2}Z_{jk}^{(2)} = \int_0^t \frac{1}{\sqrt{n}} W_{jk}(s) \left( \frac{1}{Y_k(s)} - \frac{1}{Y_\bullet(s)} \right) dM_{jk}(s) - \int_0^t \frac{1}{\sqrt{n}} W_{jk}(s) \frac{1}{Y_\bullet(s)} \sum_{h \neq k} dM_{jh}(s)
\]

(3.26)
Let \( H_1(t) = \frac{1}{\sqrt{n}} W_{jk}(s) \left( \frac{1}{Y_k(s)} - \frac{1}{Y_\ast(s)} \right) \), \( H_2(t) = -\frac{1}{\sqrt{n}} W_{jk}(s) \frac{1}{Y_\ast(s)} \), \( \lambda_1(t) = Y_k(t) h_{j0}(t) \), and \( \lambda_2(t) = Y_\ast(t) h_{j0}(t) \). To satisfy the first condition of the martingale central limit theorem, the quantity \( H_1(t)^2 \lambda_1(t) + H_2(t) \lambda_2(t) \) must converge to a deterministic function:

\[
H_1(t)^2 \lambda_1(t) + H_2(t) \lambda_2(t) = \frac{(W_{jk}(t))^2}{n} \left( \frac{1}{Y_k(s)} - \frac{1}{Y_\ast(s)} \right) h_{j0}(s) ds \overset{p}{\to} (W_{jk}^0)^2 \left( \frac{1}{y_k} - \frac{1}{y_\ast} \right) h_{j0}(s) ds
\]

Thus, the variance of \( V_{jk}^{(2)}(t) \) is defined as

\[
\text{Var}[V_{jk}^{(2)}(t)] = \int_0^t (W_{jk}^0)^2 \left( \frac{1}{y_k} - \frac{1}{y_\ast} \right) h_{j0}(s) ds \quad (3.27)
\]

For the case of \( k \neq k' \):

\[
\text{Cov}[V_{jk}^{(2)}(t), V_{jk'}^{(2)}(t)] = -\int_0^t W_{jk'}^0 \frac{1}{y_\ast} h_{j0}(s) ds \quad (3.28)
\]

Hence, the asymptotic variance-covariances of \( V_{jk}^{(2)}(t) \) are given by

\[
\text{Cov}[V_{jk}^{(2)}(t), V_{jk'}^{(2)}(t)] = \int_0^t W_{jk}^0 W_{jk'}^0 \left( \frac{1}{y_k} - \frac{1}{y_\ast} \right) h_{j0}(s) ds \quad (3.29)
\]

Hence, each test process is asymptotically distributed as a Gaussian mar-
tingale with a specific variance-covariance process. These asymptotic variance-
covariance processes can be estimated by

\[
\hat{\sigma}_{jk,j'k'} = \frac{1}{n} \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\gamma_{jj'}\gamma_{kk'}}{Y_k(s)} - \frac{1}{JY_\bullet(s)} \right) \frac{dN_\bullet(s)}{JY_\bullet(s)},
\]

\[
\hat{\sigma}_{jk,j'k'} = \frac{1}{n} \int_0^t W_{jk}(s)W_{j'k'}(s) \left( \frac{\delta_{kk'}}{Y_k(s)} - \frac{1}{Y_\bullet(s)} \right) \frac{dN_\bullet(s)}{Y_k(s)},
\]

\[
\hat{\sigma}_{jk,j'k'} = \frac{1}{n} \int_0^t W_{jk}(s)W_{j'k}(s) \left( \frac{\delta_{j,j'}}{Y_k(s)} - \frac{1}{JY_k(s)} \right) \frac{dN_{j\bullet}(s)}{Y_k(s)}.
\]

The test processes for each \( j \) and \( k \) are incorporated into the following matrices:

\[
Z^{(1)} = \left[ Z_{11}^{(1)}(t), \ldots, Z_{1K}^{(1)}(t), Z_{21}^{(1)}(t), \ldots, Z_{2K}^{(1)}(t), \ldots, Z_{JK}^{(1)}(t) \right]',
\]

\[
Z^{(2)} = \left[ Z_{j0}^{(2)}, \ldots, Z_{j0}^{(2)} \right]',
\]

\[
Z^{(3)} = \left[ Z_{j0}^{(3)}, \ldots, Z_{0K}^{(3)} \right]',
\]

where \( Z_{j0}^{(2)} = \left[ Z_{j1}^{(2)}(t), \ldots, Z_{jK}^{(2)}(t) \right]' \) and \( Z_{0k}^{(3)} = \left[ Z_{1k}^{(3)}(t), \ldots, Z_{jk}^{(3)}(t) \right]' \). The estimated variance-covariance matrices for these vectors of test processes are defined as the following:

\[
\mathbf{\hat{\Sigma}}^{(1)} = \left[ \hat{\sigma}_{jk,j'k'} \right]_{JK \times JK},
\]

\[
\mathbf{\hat{\Sigma}}^{(2)} = \text{diag} \left[ \mathbf{\hat{\Sigma}}_{j0}^{(2)}, \ldots, \mathbf{\hat{\Sigma}}_{j0}^{(2)} \right],
\]

\[
\mathbf{\hat{\Sigma}}^{(3)} = \text{diag} \left[ \mathbf{\hat{\Sigma}}_{01}^{(3)}, \ldots, \mathbf{\hat{\Sigma}}_{0K}^{(3)} \right],
\]
where \( \tilde{\Sigma}_{j0}^{(2)} = [(\hat{\sigma}_{j,kk'})]_{K \times K} \) and \( \tilde{\Sigma}_{0k}^{(3)} = [(\hat{\sigma}_{j,k})]_{J \times J} \).

The summation of the test processes \( \sum Z^{(1)}_{jk}, \sum Z^{(2)}_{jk}, \sum Z^{(3)}_{jk} \) are all equal to 0. This means the test process matrices formed by deleting their last entry provide the same information as the test process matrices with all entries. Specifically, the element \( Z^{(1)}_{jK}(t) \) can be deleted from the matrix \( Z^{(1)} \), the element \( Z^{(2)}_{jK}(t) \) can be deleted from each vector \( Z^{(2)}_{j0} \) in the matrix \( Z^{(2)} \), and the element \( Z^{(3)}_{jK}(t) \) can be deleted from each vector \( Z^{(3)}_{0k} \) in the matrix \( Z^{(3)} \). The test process matrices used in the test statistics are the following:

\[
Z^{*(1)} = \left[ Z_{11}^{(1)}(t), \ldots, Z_{1K}^{(1)}(t), Z_{21}^{(1)}(t), \ldots, Z_{2K}^{(1)}(t), \ldots, Z_{JK}^{(1)}(t) \right]',
\]

\[
Z^{*(2)} = \left[ Z_{10}^{*(2)}, \ldots, Z_{j0}^{*(2)} \right]',
\]

\[
Z^{*(3)} = \left[ Z_{01}^{*(3)}, \ldots, Z_{0K}^{*(3)} \right]',
\]

where \( Z^{*(2)}_{j0} = \left[ Z_{j1}^{(2)}(t), \ldots, Z_{JK-1}^{(2)}(t) \right]' \) and \( Z^{*(3)}_{0k} = \left[ Z_{1k}^{(3)}(t), \ldots, Z_{J-1k}^{(3)}(t) \right]' \).

The variance-covariance matrices \( \tilde{\Sigma}^{*^{(1)}}, \tilde{\Sigma}^{*^{(2)}}, \tilde{\Sigma}^{*^{(3)}} \) used in the test statistics are formed by deleting their last row and column. The test statistics for each hypothesis are defined as

\[
T_1 = \frac{1}{n} Z^{*(1)' \left( \tilde{\Sigma}^{*^{(1)}} \right)^{-1} Z^{*(1)}},
\]

\[
T_2 = \frac{1}{n} Z^{*(2)' \left( \tilde{\Sigma}^{*^{(2)}} \right)^{-1} Z^{*(2)}},
\]

\[
T_3 = \frac{1}{n} Z^{*(3)' \left( \tilde{\Sigma}^{*^{(3)}} \right)^{-1} Z^{*(3)}},
\]
The test statistics are standardizations of the test process matrices, hence the test statistics $T_1, T_2,$ and $T_3$ converge to $\chi^2_{JK-1}, \chi^2_{J(K-1)},$ and $\chi^2_{(J-1)K}$ random variables, respectively.

3.4.3 Weight Functions

Different weight functions will be analyzed in the analysis:

\[ W^0_{jk} = Y_k(t) \]
\[ W^1_{jk} = Y_k(t) \left( \frac{Y(t)}{n} \right)^\rho \text{ where } \rho \geq 0 \]
\[ W^2_{jk} = Y_k(t) (S_k(t) - \hat{S}_k(t))^\rho (1 - \hat{S}_k(t))^r \text{ where } \rho \geq 0 \text{ and } r \geq 0 \]
\[ W^3_{jk} = Y_k(t) (1 - \hat{F}_j(t))^\rho \text{ where } 0 < \rho < 1 \]

When $\rho = 1$ and $r = 0$, the weight function $W^2_{jk}$ becomes the Kaplan-Meier estimate for the survival function, treating all competing risks as one type of event. More emphasis is placed on the early differences between the competing risks, with events near the beginning of the time period assigned more weight than later events. Conversely, the weight function $W^3_{jk}$ becomes the estimate for the cumulative incidence function for $\rho = 0$ and $r = 1$, which places heavier weight on the later event times than the earlier event times.

Through simulation studies, the authors established that the test statistics using the weight function $W^1_{jk}$ are accurate for small sample sizes and the power
of each test increases when the event times within each competing risk are highly correlated. This research will analyze the additional weight functions as well as different levels of censoring to examine the performance of the test statistics under these situations.
CHAPTER IV

MONTE CARLO SIMULATION STUDY

4.1 Simulation Procedures

Tiwari et al (2006) showed that using the default weight function $W_{jk}^0 = Y_k(t)$ given in 3.30 in the three subtests controlled type I error and demonstrated increased power over parametric tests using multivariate normal and exponential distributions to simulate competing risks data. In order to evaluate the performance in terms of type I error rates and power for the other weight functions $W_{jk}^1, W_{jk}^2, W_{jk}^3$ given in 3.30 in the three TKC subtests detailed in section 3.4, a Monte Carlo simulation study will be conducted. We first set up the notation. Let $T_{ik}^j$ denote the random variable corresponding to the discontinuation time of the $i$th subject in the $k$th group due to the $j$th category of discontinuation. Let $\delta_{ik}$ denote the random variable corresponding to the type of discontinuation observed for the $i$th subject in the $k$th group, with $\delta_{ik} = 0$ for subjects completing the study. Then the bivariate random variable pair $(X_{ik}^0, \delta_{ik}^0)$ is given by

\[
X_{ik}^0 = \min(T_{ik}^1, \ldots, T_{ik}^J)
\]

\[
\delta_{ik}^0 = \text{index } j \text{ minimizing } T_{ik}^j, \text{ where } j = 1, \ldots, J
\]
The simulation data were generated from a Weibull distribution with shape parameter $\theta$ and scale parameter $\beta$:

$$f(T_{ik}^j; \theta, \beta) = \frac{\theta}{\beta} \left( \frac{T_{ik}^j}{\beta} \right)^{\theta-1} \exp \left( - \left( \frac{T_{ik}^j}{\beta} \right)^{\theta} \right), 0 \leq T_{ik}^j < \infty, \theta > 0, \beta > 0 \quad (4.1)$$

The Weibull distribution has the advantage of modelling either increasing or decreasing risks of discontinuation by defining the shape parameter $\theta$ to be less than 1 or greater than 1, respectively (Lee & Go, 1997). The Weibull distribution has a broader application for simulating failure time data compared to the exponential distribution. Assume that the scale parameter $\beta = 1$. Denote the discontinuation time vector for group $k$ as $T_k = (T_{ik1}, \ldots, T_{ikJ})$, and let the shape parameter vector for group $k$ be $\theta_k = (\theta_{ik1}, \ldots, \theta_{ijk})'$. The particular value of the shape parameters associated with the discontinuation time $T_{ik}^j$ in each group will vary depending on which TKC subtest is analyzed.

Let the shape parameter matrix be $\Theta = (\theta_1, \ldots, \theta_K)$, with columns for each group and rows for each risk of discontinuation. To generate data under the null hypotheses for each of the tests in the TKC procedure, the following cases for $\Theta$ are defined:

- Data generation under $H_{01}$: $\theta_k = (\theta_{00}, \ldots, \theta_{00})'$. This gives each discontinuation time the same shape parameter $\theta_{00}$ for its distribution, no matter
which risk of discontinuation or group the time represents.

\[ \Theta_1 = (\theta_1, \ldots, \theta_K) = \begin{bmatrix} \theta_{00} & \theta_{00} & \cdots & \theta_{00} \\ \theta_{00} & \theta_{00} & \cdots & \theta_{00} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{00} & \theta_{00} & \cdots & \theta_{00} \end{bmatrix} \]

- Data generation under \( H_{02} \): \( \theta_k = (\theta_{10}, \ldots, \theta_{J0})' \). The following matrix gives each risk of discontinuation the same \( \theta \) values, but different \( \theta \) values for each group.

\[ \Theta_2 = (\theta_1, \ldots, \theta_K) = \begin{bmatrix} \theta_{10} & \theta_{10} & \cdots & \theta_{10} \\ \theta_{20} & \theta_{20} & \cdots & \theta_{20} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{J0} & \theta_{J0} & \cdots & \theta_{J0} \end{bmatrix} \]

- Data generation under \( H_{03} \): \( \theta_k = (\theta_{0k}, \ldots, \theta_{0k})' \). The following matrix gives each group the same \( \theta \) values, but different \( \theta \) values for each risk of discontinuation.

\[ \Theta_3 = (\theta_1, \ldots, \theta_K) = \begin{bmatrix} \theta_{01} & \theta_{02} & \cdots & \theta_{0K} \\ \theta_{01} & \theta_{02} & \cdots & \theta_{0K} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{01} & \theta_{02} & \cdots & \theta_{0K} \end{bmatrix} \]

To generate data under the alternatives to each null hypothesis, particular
elements of the matrix $\Theta$ are changed in order to create a matrix that is different than the corresponding matrix under the null hypothesis. Suppose $J = 3$ and $K = 3$. Then the following matrices are defined under the alternatives:

$$\Theta_1(\gamma) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \gamma & 1 \\ 1 & 1 & 1 \end{bmatrix}, \Theta_2(\gamma) = \begin{bmatrix} 0.5 & 0.5 & 0.5 \\ 1 & \gamma & 1 \\ 2 & 2 & 2 \end{bmatrix}, \Theta_3(\gamma) = \begin{bmatrix} 0.5 & 1 & 2 \\ 0.5 & \gamma & 2 \\ 0.5 & 1 & 2 \end{bmatrix}.$$  

where $\Theta_1(\gamma = 1) = \Theta_1$ under $H_{01}$, $\Theta_2(\gamma = 1) = \Theta_2$ under $H_{02}$, and $\Theta_3(\gamma = 1) = \Theta_3$ under $H_{03}$. Censoring times were generated using the exponential distribution with rate $\beta$, where $\beta$ is determined by prespecified censoring proportions:

$$f(x; \beta) = \beta e^{-\beta x}, 0 \leq x < \infty, \beta > 0. \quad (4.2)$$

The size of the tests is estimated based on 1000 replicates with sample sizes of 200 for each of the three groups. For each data set, the null hypothesis was tested at the 5% level. The number of times each test rejected the null hypothesis was counted and the size of the test was estimated using the following:

$$\hat{\alpha} = \frac{\text{Number of rejections}}{1,000} \times 100\% \quad (4.3)$$

Table 1 contains the configuration for the simulations examining the type I error rate of the three TKC subtests using the different weight functions. This
configuration was used for each of the three TKC subtests. Different censoring conditions were produced for each weight function implementation. By increasing the $\beta$ parameter, increasing proportions of censored data can be generated. Light censoring (3-14 percent) of the data were generated using $\beta = 0.1$ and $\beta = 0.5$, moderate censoring (25-40 percent) of the data were generated using $\beta = 1$ and $\beta = 2$, and heavy censoring (87 percent) of the data were generated using $\beta = 20$ in the censoring distribution given in 4.2.

Table 1: Configurations of TKC Type I Error Rate Simulations

<table>
<thead>
<tr>
<th>Weight Function</th>
<th>$\rho$</th>
<th>$r$</th>
<th>$\gamma$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{jk}^1 = Y_k(t) \left( \frac{Y_n(t)}{n} \right)^\rho$</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>0.1,0.5,1,2,20</td>
</tr>
<tr>
<td>$W_{jk}^1 = Y_k(t) \left( \frac{Y_n(t)}{n} \right)^\rho$</td>
<td>0</td>
<td>NA</td>
<td>1</td>
<td>0.1,0.5,1,2,20</td>
</tr>
<tr>
<td>$W_{jk}^2 = Y_k(t)(S_k(t))^\rho(1 - \tilde{S}_k(t))^r$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.1,0.5,1,2,20</td>
</tr>
<tr>
<td>$W_{jk}^2 = Y_k(t)(\tilde{S}_k(t))^\rho(1 - \tilde{S}_k(t))^r$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.1,0.5,1,2,20</td>
</tr>
<tr>
<td>$W_{jk}^3 = Y_k(t)(1 - F_j(t))^\rho$</td>
<td>1</td>
<td>NA</td>
<td>1</td>
<td>0.1,0.5,1,2,20</td>
</tr>
<tr>
<td>$W_{jk}^3 = Y_k(t)(1 - F_j(t))^\rho$</td>
<td>0</td>
<td>NA</td>
<td>1</td>
<td>0.1,0.5,1,2,20</td>
</tr>
</tbody>
</table>

Table 2 contains the configuration for the simulations examining the power of the three TKC subtests using the different weight functions. This configuration was used for each of the three TKC subtests. Departures from the null hypothesis for each TKC subtest were controlled by varying the $\gamma$ parameters in the $\Theta_1(\gamma), \Theta_2(\gamma), \Theta_3(\gamma)$ matrices. Moderate censoring (25-40 percent) of the data were generated using $\beta = 1$ and heavy censoring (65-78 percent) of the data were generated using $\beta = 10$ in the censoring distribution given in 4.2.
Table 2: Configurations of TKC Power Simulations

<table>
<thead>
<tr>
<th>Weight Function</th>
<th>$\rho$</th>
<th>$r$</th>
<th>$\gamma$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{jk}^1 = Y_k(t) \left( \frac{Y_k(t)}{n} \right)^{\rho}$</td>
<td>1</td>
<td>NA</td>
<td>0.5,1,2,4,8</td>
<td>1, 10</td>
</tr>
<tr>
<td>$W_{jk}^1 = Y_k(t) \left( \frac{Y_k(t)}{n} \right)^{\rho}$</td>
<td>0</td>
<td>NA</td>
<td>0.5,1,2,4,8</td>
<td>1, 10</td>
</tr>
<tr>
<td>$W_{jk}^2 = Y_k(t) (\hat{S}_k(t))^{\rho}(1 - \hat{S}_k(t))^{r}$</td>
<td>1</td>
<td>0</td>
<td>0.5,1,2,4,8</td>
<td>1, 10</td>
</tr>
<tr>
<td>$W_{jk}^2 = Y_k(t) (\hat{S}_k(t))^{\rho}(1 - \hat{S}_k(t))^{r}$</td>
<td>0</td>
<td>1</td>
<td>0.5,1,2,4,8</td>
<td>1, 10</td>
</tr>
<tr>
<td>$W_{jk}^3 = Y_k(t) (1 - \hat{F}_j(t))^{\rho}$</td>
<td>1</td>
<td>NA</td>
<td>0.5,1,2,4,8</td>
<td>1, 10</td>
</tr>
<tr>
<td>$W_{jk}^3 = Y_k(t) (1 - \hat{F}_j(t))^{\rho}$</td>
<td>0</td>
<td>NA</td>
<td>0.5,1,2,4,8</td>
<td>1, 10</td>
</tr>
</tbody>
</table>

4.2 Simulation Results

4.2.1 Introduction

As described in section 4.1, data were simulated for $J = 3$ competing risks and $K = 3$ groups with three different configurations of the shape parameter matrices $\Theta_1(\gamma), \Theta_2(\gamma)$ and $\Theta_3(\gamma)$ corresponding to the three different TKC tests. The sample size for each of the $K = 3$ groups was 200, and the proportion of censored observations varied according to the censoring parameter $\beta$. For each shape parameter $\gamma$, weight function parameters $\rho$ and $r$, and censoring parameter $\beta$ configuration, 1000 datasets were simulated and summarized. The simulations were conducted using the R statistical software package (R Development Core Team, 2008), and the simulation programs can be found in listings A, A, and A of the Appendix.
4.2.2 Impact of Censoring Results

Table 3 contains the empirical type I error estimates using the different censoring conditions for all three TKC tests. In the case of $\rho = 1$, the range of the empirical type I error rates for all three TKC tests was 2.9 to 6.8 percent. This range is in agreement with the empirical type I error rates using $W^0_{jk}$, which are the weight functions $W^1_{jk}$ and $W^3_{jk}$ with $\rho = 0$. The amount of censored data does not seem to have a substantial effect on the empirical type I error rates. For example, using $W^3_{jk}$ and $\rho = 0$, the empirical type I error rates range from 3.6 to 6.6 percent in all three TKC test procedures. Considering the second weight function $W^2_{jk}$, one outlier of the empirical type one error rate of 17.4 percent was present for the configuration of $W^2_{jk}, \beta = 20, \rho = 0$ in TKC test 1. However, the range of empirical type one error rates for $W^2_{jk}$ and $\rho = 1$ was from 2.5 to 6.6 percent.

4.2.3 Impact of Weight Functions Results

Table 4 contains the empirical power estimates using the different weight functions for all three TKC tests. Using $W^1_{jk}$ with $\rho = 1$ increased the empirical power of all three TKC tests from 8 to 33 percent in light censoring conditions ($\beta = 1$) over the default weight function with $\rho = 0$. In the case of moderate censoring ($\beta = 10$), using $W^1_{jk}$ with $\rho = 1$ decreased the empirical power when the hazards of DC were increasing ($\gamma > 1$) in TKC tests 1 and 2. Using $W^2_{jk}$ with
Table 3: Empirical Type I Error Rates Under Different Censoring Conditions

<table>
<thead>
<tr>
<th>Weight</th>
<th>$\beta$</th>
<th>TKC Test 1</th>
<th>TKC Test 2</th>
<th>TKC Test 3</th>
<th>TKC Test 1</th>
<th>TKC Test 2</th>
<th>TKC Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W^1_{jk}$</td>
<td>0.1</td>
<td>0.06</td>
<td>0.053</td>
<td>0.049</td>
<td>0.055</td>
<td>0.049</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.056</td>
<td>0.052</td>
<td>0.048</td>
<td>0.041</td>
<td>0.043</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.047</td>
<td>0.057</td>
<td>0.051</td>
<td>0.059</td>
<td>0.052</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.057</td>
<td>0.061</td>
<td>0.038</td>
<td>0.047</td>
<td>0.054</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.045</td>
<td>0.037</td>
<td>0.029</td>
<td>0.04</td>
<td>0.043</td>
<td>0.029</td>
</tr>
<tr>
<td>$W^2_{jk}$</td>
<td>0.1</td>
<td>0.059</td>
<td>0.047</td>
<td>0.059</td>
<td>0.058</td>
<td>0.052</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.062</td>
<td>0.066</td>
<td>0.061</td>
<td>0.054</td>
<td>0.066</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.065</td>
<td>0.051</td>
<td>0.05</td>
<td>0.054</td>
<td>0.035</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.057</td>
<td>0.054</td>
<td>0.059</td>
<td>0.062</td>
<td>0.054</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.174</td>
<td>0.053</td>
<td>0.029</td>
<td>0.051</td>
<td>0.04</td>
<td>0.025</td>
</tr>
<tr>
<td>$W^3_{jk}$</td>
<td>0.1</td>
<td>0.048</td>
<td>0.049</td>
<td>0.055</td>
<td>0.048</td>
<td>0.04</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.039</td>
<td>0.046</td>
<td>0.039</td>
<td>0.047</td>
<td>0.039</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.052</td>
<td>0.051</td>
<td>0.048</td>
<td>0.054</td>
<td>0.053</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.059</td>
<td>0.055</td>
<td>0.06</td>
<td>0.05</td>
<td>0.045</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.048</td>
<td>0.028</td>
<td>0.036</td>
<td>0.047</td>
<td>0.038</td>
<td>0.041</td>
</tr>
</tbody>
</table>

1. $W^1_{jk}(t) = Y_k(t)\left(\frac{Y_n(t)}{n}\right)^\rho$
2. $W^2_{jk}(t) = Y_k(t)(\hat{S}_k(t))^{\rho}(1 - \hat{S}_k(t))^r$
3. $W^3_{jk}(t) = Y_k(t)(1 - \hat{F}_j(t))^\rho$

$\rho = 1$ and $r = 0$ produced substantially higher empirical power in all conditions compared to setting $\rho = 0$ and $r = 1$. Using $W^3_{jk}$ with $\rho = 1$ in all three TKC tests produced little to no increases in empirical power over the default weight function.

Overall, in the case of moderate censoring ($\beta = 10$) the weight functions $W^2_{jk}$ and $W^3_{jk}$ only increased the empirical power by 2 to 4 percent over the default weight function, and $W^1_{jk}$ actually decreased the empirical power in the case of increasing hazards of DC.
Table 4: Empirical Power for TKC Tests

<table>
<thead>
<tr>
<th>Weight $W_{jk}^i$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>$\rho = 0$ TKC Test 1</th>
<th>$\rho = 0$ TKC Test 2</th>
<th>$\rho = 0$ TKC Test 3</th>
<th>$\rho = 1$ TKC Test 1</th>
<th>$\rho = 1$ TKC Test 2</th>
<th>$\rho = 1$ TKC Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W_{jk}^1$</td>
<td>0.5</td>
<td>1</td>
<td>0.918</td>
<td>0.906</td>
<td>0.906</td>
<td>0.999</td>
<td>0.993</td>
<td>0.995</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.806</td>
<td>0.649</td>
<td>0.737</td>
<td>0.998</td>
<td>0.97</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>1</td>
<td>0.992</td>
<td>0.998</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$W_{jk}^2$</td>
<td>0.5</td>
<td>1</td>
<td>0.999</td>
<td>0.979</td>
<td>0.979</td>
<td>1</td>
<td>0.996</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.78</td>
<td>0.698</td>
<td>0.768</td>
<td>0.723</td>
<td>0.653</td>
<td>0.771</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.989</td>
<td>0.963</td>
<td>0.989</td>
<td>0.866</td>
<td>0.832</td>
<td>0.937</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.999</td>
<td>0.981</td>
<td>0.99</td>
<td>0.867</td>
<td>0.836</td>
<td>0.944</td>
</tr>
<tr>
<td>$W_{jk}^3$</td>
<td>0.5</td>
<td>1</td>
<td>0.131</td>
<td>0.136</td>
<td>0.076</td>
<td>0.999</td>
<td>0.988</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.128</td>
<td>0.138</td>
<td>0.084</td>
<td>0.996</td>
<td>0.964</td>
<td>0.977</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.412</td>
<td>0.461</td>
<td>0.308</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.69</td>
<td>0.699</td>
<td>0.547</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$W_{jk}^3$</td>
<td>0.5</td>
<td>1</td>
<td>0.689</td>
<td>0.617</td>
<td>0.393</td>
<td>1</td>
<td>0.996</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.389</td>
<td>0.406</td>
<td>0.235</td>
<td>0.802</td>
<td>0.731</td>
<td>0.809</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.702</td>
<td>0.857</td>
<td>0.629</td>
<td>0.993</td>
<td>0.969</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.786</td>
<td>0.9</td>
<td>0.708</td>
<td>0.999</td>
<td>0.976</td>
<td>0.999</td>
</tr>
<tr>
<td>$W_{jk}^3$</td>
<td>0.5</td>
<td>1</td>
<td>0.898</td>
<td>0.738</td>
<td>0.719</td>
<td>0.935</td>
<td>0.803</td>
<td>0.753</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.831</td>
<td>0.67</td>
<td>0.677</td>
<td>0.853</td>
<td>0.68</td>
<td>0.672</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.999</td>
<td>0.981</td>
<td>0.996</td>
<td>1</td>
<td>0.984</td>
<td>0.997</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>1</td>
<td>0.997</td>
<td>0.999</td>
<td>1</td>
<td>0.995</td>
<td>1</td>
</tr>
<tr>
<td>$W_{jk}^3$</td>
<td>0.5</td>
<td>1</td>
<td>0.997</td>
<td>0.98</td>
<td>0.974</td>
<td>0.997</td>
<td>0.985</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>0.755</td>
<td>0.683</td>
<td>0.74</td>
<td>0.789</td>
<td>0.703</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>0.987</td>
<td>0.959</td>
<td>0.992</td>
<td>0.992</td>
<td>0.962</td>
<td>0.992</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>0.997</td>
<td>0.973</td>
<td>0.999</td>
<td>0.997</td>
<td>0.967</td>
<td>0.999</td>
</tr>
</tbody>
</table>

1 $W_{jk}^1(t) = Y_k(t) \left( \frac{Y_k(t)}{n} \right)^\beta$

2 $W_{jk}^2(t) = Y_k(t)(\hat{S}_k(t))(1 - \hat{S}_k(t))^\gamma$

3 $W_{jk}^3(t) = Y_k(t)(1 - \hat{F}_j(t))^\rho$
CHAPTER V

APPLICATION OF COMPETING RISKS METHODS

5.1 Introduction

All analyses were conducted using the R statistical software package (R Development Core Team, 2008). Specific routines written to conduct the TKC and KLY tests can be found in listings A and A of the Appendix. The default weight functions $w(t) = Y_k(t)$ and $W_{jk}^0 = Y_k(t)$ for the KLY and TKC tests will be used in all applications unless stated otherwise. Cox regression models were generated using the survival package (Therneau & original R port by Thomas Lumley, 2009), and Fine & Gray regression models were generated using the cmprsk package (Gray, 2008).

5.2 Baseline Results

Table 5 contains the number of patients who discontinued due to various reasons and patients who completed the study. In the DPNP, FM, and GAD disease states, discontinuations due to adverse events were higher than discontinuations due to lack of efficacy, while for MDD clinical trials the opposite was true.

Table 6 contains summary statistics for the patients in each disease state.
Table 5: Number (percent) of Patients who DC or Complete

<table>
<thead>
<tr>
<th></th>
<th>Complete</th>
<th>DC=AE</th>
<th>DC=LOE</th>
<th>DC=OTH</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPNP</td>
<td>887 (77.9)</td>
<td>118 (10.4)</td>
<td>19 (1.7)</td>
<td>115 (10.1)</td>
<td>1139</td>
</tr>
<tr>
<td>FM</td>
<td>822 (58.3)</td>
<td>234 (16.6)</td>
<td>133 (9.4)</td>
<td>222 (15.7)</td>
<td>1411</td>
</tr>
<tr>
<td>GAD</td>
<td>1276 (66.9)</td>
<td>199 (10.4)</td>
<td>78 (4.1)</td>
<td>355 (18.6)</td>
<td>1908</td>
</tr>
<tr>
<td>MDD</td>
<td>2146 (65.6)</td>
<td>274 (8.4)</td>
<td>304 (9.3)</td>
<td>546 (16.7)</td>
<td>3270</td>
</tr>
</tbody>
</table>

Patients in MDD and GAD studies had average ages of approximately 42 years.

The majority of patients in each disease state were Caucasian (78%-88%). In the GAD, FM, and MDD disease states, the majority of patients were female (95%, 62%, and 64%, respectively).

Table 6: Descriptive Statistics by Disease State

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>DPNP</th>
<th>FIBRO</th>
<th>GAD</th>
<th>MDD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N = 1139</td>
<td>N = 1411</td>
<td>N = 1908</td>
<td>N = 3270</td>
</tr>
<tr>
<td>Age</td>
<td>7728</td>
<td>(59.90±10.62)</td>
<td>(50.24±10.98)</td>
<td>(42.33±13.33)</td>
<td>(41.59±12.37)</td>
</tr>
<tr>
<td>Race : CAU</td>
<td>7657</td>
<td>85% (961)</td>
<td>88% (1234)</td>
<td>78% (1494)</td>
<td>85% (2720)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4% (48)</td>
<td>2% (33)</td>
<td>7% (127)</td>
<td>7% (236)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9% (105)</td>
<td>9% (127)</td>
<td>11% (204)</td>
<td>7% (212)</td>
</tr>
<tr>
<td>AF</td>
<td></td>
<td>0% (0)</td>
<td>0% (3)</td>
<td>0% (2)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>HIS</td>
<td></td>
<td>2% (17)</td>
<td>1% (9)</td>
<td>4% (80)</td>
<td>1% (45)</td>
</tr>
<tr>
<td>NA</td>
<td></td>
<td>0% (0)</td>
<td>0% (3)</td>
<td>0% (2)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>ASIAN</td>
<td></td>
<td>57% (647)</td>
<td>5% (73)</td>
<td>38% (721)</td>
<td>36% (1171)</td>
</tr>
</tbody>
</table>

\( x \pm s \) represents \( \bar{X} \pm 1 \) SD.

\( N \) is the number of non-missing values.

Numbers after percents are frequencies.
5.3 Competing Risks Testing Results

The GAD and MDD disease states will be used to demonstrate the competing risks analyses of discontinuation. Figure 1 contains the plots of the estimated cumulative incidence functions for each disease state. In the GAD disease states shown in Figure 1(a), the CIF for DC due to AE is consistently higher than the CIF for DC due to LOE in the duration of the study periods. In the MDD disease state, Figure 1(b) shows that the CIF curves for DC due to AE and DC due to LOE cross somewhere between week 6 and week 8 of the study period. The CIF for DC due to AE is higher before this point, and lower than DC due to LOE after this point.

![Cumulative Incidence Curves for GAD](image1)

(a) Cumulative Incidence Curves for GAD

![Cumulative Incidence Curves for MDD](image2)

(b) Cumulative Incidence Curves for MDD

Figure 1: Cumulative Incidence Curves

Table 7 contains the results of conducting the TKC test 3 and KLY tests on the competing risks of DC due to AE and DC due to LOE, with treating DC
due to other reasons as censored observations and using the weight functions of
\( w(t) = Y_k(t) \) and \( W_{jk}^0 = Y_k(t) \) for the KLY and TKC tests, respectively. For
the disease state GAD, both the TKC test 3 and the KLY tests concluded that
the risks of discontinuation were significantly different. However, the KLY test
concluded that the risks of DC due to AE and DC due to LOE were significantly
different in MDD studies, while the TKC test 3 did not determine a significant
difference. One possible reason for this result is the way the test statistics for
the KLY and TKC tests are constructed. The KLY test uses the standardized
maximum difference between the CIF’s of the two competing risks, while the TKC
test statistic is a summation of the deviations for each competing risks’ CSH from
the estimated value of the CSH under the null hypothesis.

Table 7: TKC Test 3 and KLY Test Results

<table>
<thead>
<tr>
<th></th>
<th>KLY Test</th>
<th>TKC Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( C_n^* ) P-value</td>
<td>( \chi_1^2 ) P-value</td>
</tr>
<tr>
<td>GAD</td>
<td>7.27 &lt;0.0001</td>
<td>52.86 &lt;0.0001</td>
</tr>
<tr>
<td>MDD</td>
<td>4.62 &lt;0.0001</td>
<td>1.56 0.2121</td>
</tr>
</tbody>
</table>

Table 8 contains the results of conducting the three TKC tests using the
different weight functions with treatment assignment as the grouping factor of
interest. The test one statistic \( T_1 \) is highly significant, indicating that the six
CCSH curves are significantly different. Figure 2 contains the estimated CCSH
curves of the categories of DC for both the placebo and treatment groups.

The plots in figures 2(a) - 2(b) show the overall difference in the risks of DC
Table 8: TKC Test Statistics for Risks of DC with Treatment Assignment

| Disease State | Weight Function | TKC Test 1 |  | TKC Test 2 |  | TKC Test 3 |  |
|---------------|-----------------|------------|----------------|------------|----------------|----------------|
|               |                 | TKC Test 1 | Test Statistic | P-value    | Test Statistic | P-value    | Test Statistic | P-value |
| GAD           | $W^1_{jk}$      | 222.77     | < 0.0001       | 69.31      | < 0.0001       | 222.72     | < 0.0001       |
|               | $W^1_{jk} (\rho = 1)$ | 207.9     | < 0.0001       | 67.72      | < 0.0001       | 206.22     | < 0.0001       |
|               | $W^2_{jk} (\rho = 1, r = 0)$ | 217.98     | < 0.0001       | 68.24      | < 0.0001       | 216.27     | < 0.0001       |
|               | $W^3_{jk} (\rho = 1)$ | 222.48     | < 0.0001       | 69.47      | < 0.0001       | 222.06     | < 0.0001       |
| MDD           | $W^1_{jk}$      | 161.03     | < 0.0001       | 56.68      | < 0.0001       | 160.88     | < 0.0001       |
|               | $W^1_{jk} (\rho = 1)$ | 149.47     | < 0.0001       | 58.86      | < 0.0001       | 149.43     | < 0.0001       |
|               | $W^2_{jk} (\rho = 1, r = 0)$ | 159.3     | < 0.0001       | 57.92      | < 0.0001       | 158.5      | < 0.0001       |
|               | $W^3_{jk} (\rho = 1)$ | 159.5     | < 0.0001       | 57.59      | < 0.0001       | 162.97     | < 0.0001       |

(a) GAD All DC Categories

(b) MDD All DC Categories

Figure 2: CCSH Curves for DC Categories and Treatment Groups

category and treatment assignment pairs. The test statistics $T_2$ and $T_3$ for TKC tests 2 and 3 are also significant. Thus at least one group of patients is significantly different from the other group within at least one of the DC categories, and at least one DC category’s CCSH is significantly different from the other DC categories within at least one of the groups of patients. To see which specific risks of DC have significant treatment assignment differences, and which specific treatment
groups have significantly different risks of DC, partitions of the $T_2$ and $T_3$ test statistics can be compared to a $\chi^2$ random variable with appropriate degrees of freedom.

Table 9 contains the partitions of the $T_2$ test statistic for the different categories of DC. Within DC due to AE and DC due to LOE, the risks of discontinuation are significantly different between the placebo and treatment groups in each disease state. However, within DC due to other reasons the placebo and treatment groups do not have significantly different risks of discontinuation.

Table 9: TKC Test 2 Partitions for Risks of DC with Treatment Assignment

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Weight Function</th>
<th>DC due to AE Test Statistic</th>
<th>DC due to LOE Test Statistic</th>
<th>DC due to OTH Test Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>P-value</td>
<td>P-value</td>
<td>P-value</td>
</tr>
<tr>
<td>GAD</td>
<td>$W_{1k}^0$</td>
<td>28.33</td>
<td>&lt; 0.0001</td>
<td>0.83</td>
</tr>
<tr>
<td></td>
<td>$W_{1k}^1(\rho = 1)$</td>
<td>30.08</td>
<td>&lt; 0.0001</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>$W_{1k}^2(\rho = 1, r = 0)$</td>
<td>30.04</td>
<td>&lt; 0.0001</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>$W_{1k}^3(\rho = 1)$</td>
<td>28.93</td>
<td>&lt; 0.0001</td>
<td>0.59</td>
</tr>
<tr>
<td>MDD</td>
<td>$W_{1k}^0$</td>
<td>19.9</td>
<td>&lt; 0.0001</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td>$W_{1k}^1(\rho = 1)$</td>
<td>20.07</td>
<td>&lt; 0.0001</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>$W_{1k}^2(\rho = 1, r = 0)$</td>
<td>20.65</td>
<td>&lt; 0.0001</td>
<td>2.03</td>
</tr>
<tr>
<td></td>
<td>$W_{1k}^3(\rho = 1)$</td>
<td>19.9</td>
<td>&lt; 0.0001</td>
<td>2.56</td>
</tr>
</tbody>
</table>

Figures 3-4 show that patients in the placebo group are at higher risk of DC due to LOE, while patients in the treatment group are at higher risk of DC due to AE in both disease states.

Table 10 contains the partitions of the $T_3$ test statistic for the different treatment groups. Within the placebo group, the risks of DC are significantly different in all disease states except FM. Within the treatment group, the risks of DC are significantly different in all disease states. Figures 5-6 show that the
highest risk of DC is due to other reasons within the placebo group in the GAD disease state, while in MDD studies the risk of DC due to lack of efficacy is comparable to the risk of DC due to other reasons after the sixth week in the
Table 10: TKC Test 3 Partitions for Risks of DC with Treatment Assignment

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Weight Function</th>
<th>Placebo Group</th>
<th>Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Test Statistic</td>
<td>P-value</td>
</tr>
<tr>
<td>GAD</td>
<td>$W_{jk}^0$</td>
<td>73.4</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>65.34</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1, r = 0)$</td>
<td>70.41</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^3(\rho = 1)$</td>
<td>73.31</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>MDD</td>
<td>$W_{jk}^0$</td>
<td>52.6</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>50.89</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1, r = 0)$</td>
<td>51.86</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^3(\rho = 1)$</td>
<td>52.8</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

The placebo group.

Table 11: TKC Test 3 Partitions for Risks of DC (DC-OTH Censored)

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Weight Function</th>
<th>Placebo Group</th>
<th>Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Test Statistic</td>
<td>P-value</td>
</tr>
<tr>
<td>GAD</td>
<td>$W_{jk}^0$</td>
<td>4.44</td>
<td>0.035</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>3.39</td>
<td>0.0655</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1, r = 0)$</td>
<td>4.02</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^3(\rho = 1)$</td>
<td>4.14</td>
<td>0.042</td>
</tr>
<tr>
<td>MDD</td>
<td>$W_{jk}^0$</td>
<td>41.75</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>38.41</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1, r = 0)$</td>
<td>39.84</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^3(\rho = 1)$</td>
<td>42.01</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

Table 11 contains the partitions of the $T_3$ test statistic when treating DC due to other reasons as censored observations. Within the placebo group, no significant differences in the risks of DC due to AE and DC due to LOE were found in any of the weight functions. However, in GAD studies when the weight function $W_{jk}^1$ with $\rho = 1$ was used in calculating the test statistic, there was not
a significant difference between the risks of DC due to AE and DC due to LOE within the placebo group, while when the other weight functions were used, there were marginal significant differences between these risks at the 5% level.

Figure 5: CCSH Curves for GAD and Treatment Groups

(a) MDD: Placebo Group  (b) MDD: Treatment Group

Figure 6: CCSH Curves for MDD and Treatment Groups
Table 12 contains the results of conducting the three TKC tests using the different weight functions with study location (United States versus outside the United States) as the grouping factor of interest. The TKC test one statistic $T_1$ is highly significant, indicating that the six CCSH curves are significantly different. Figure 7 contains the estimated CCSH curves of the categories of DC for both the subjects treated in the United States and the subjects treated outside the United States in both disease states. The plots show an overall difference between the risk and study location combinations. Table 12 also contains the $T_2$ and $T_3$ test statistics, and both are significant at the 5% level. With significant differences between the two study locations within each competing risk of DC, and significant differences between the three risks of DC within a study location, partitions of the $T_2$ and $T_3$ test statistics show which particular levels of DC category or study location result in significant differences.

Table 12: TKC Test Statistics and P-values for Risks of DC with Study Location

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Weight Function</th>
<th>TKC Test 1</th>
<th>TKC Test 2</th>
<th>TKC Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>$W_{jk}$</td>
<td>246.64</td>
<td>46.38</td>
<td>209.22</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>233.92</td>
<td>44.43</td>
<td>191.68</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1, r = 0)$</td>
<td>235.71</td>
<td>45.41</td>
<td>200.72</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>235.7</td>
<td>45.86</td>
<td>203.13</td>
</tr>
<tr>
<td>MDD</td>
<td>$W_{jk}$</td>
<td>158.17</td>
<td>51.75</td>
<td>157.69</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>155.97</td>
<td>55.98</td>
<td>155.18</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1, r = 0)$</td>
<td>159.57</td>
<td>50.25</td>
<td>155.93</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}(\rho = 1)$</td>
<td>158.32</td>
<td>41.14</td>
<td>161.01</td>
</tr>
</tbody>
</table>

Table 13 contains the partitions of the $T_2$ test statistic for the different categories of DC. In the GAD and MDD disease states, risk of DC due to adverse
DC due to AE: Non-US
DC due to AE: US
DC due to LOE: Non-US
DC due to LOE: US
DC due to OTH: Non-US
DC due to OTH: US

Figure 7: CCSH Curves for DC Categories and Study Location

events did not differ significantly among the two study locations. Figures 8(b) and 9(b) show that patients treated outside of the United States were at higher risk of DC due to lack of efficacy than patients treated in the United states after four or five weeks of observation.

Table 13: TKC Test 2 Partitions for Risks of DC with Study Location

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Weight Function</th>
<th>DC due to AE</th>
<th>P-value</th>
<th>DC due to LOE</th>
<th>P-value</th>
<th>DC due to OTH</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>$W^0$</td>
<td>0.52</td>
<td>0.4707</td>
<td>11.63</td>
<td>0.0006</td>
<td>34.24</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W^1_p=1$</td>
<td>0.25</td>
<td>0.6195</td>
<td>10.59</td>
<td>0.0011</td>
<td>33.59</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W^2_p=1,r=0$</td>
<td>0.33</td>
<td>0.5632</td>
<td>11.32</td>
<td>0.0008</td>
<td>33.76</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W^3_p=1$</td>
<td>0.59</td>
<td>0.4423</td>
<td>11.86</td>
<td>0.0006</td>
<td>33.41</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>MDD</td>
<td>$W^0$</td>
<td>1.38</td>
<td>0.2394</td>
<td>41.65</td>
<td>&lt; 0.0001</td>
<td>8.72</td>
<td>0.0031</td>
</tr>
<tr>
<td></td>
<td>$W^1_p=1$</td>
<td>0.45</td>
<td>0.5013</td>
<td>38.41</td>
<td>&lt; 0.0001</td>
<td>17.11</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W^2_p=1,r=0$</td>
<td>0.66</td>
<td>0.418</td>
<td>37.06</td>
<td>&lt; 0.0001</td>
<td>12.53</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W^3_p=1$</td>
<td>0.44</td>
<td>0.5072</td>
<td>27.7</td>
<td>&lt; 0.0001</td>
<td>13.01</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Table 14 contains the partitions of the $T_3$ test statistic for the different categories of DC. Within the United States, the risks of DC are significantly
different in all disease states. Figures 10(b) and 11(b) show that the risk of DC due to LOE is higher than the other risks of DC for most of the study duration. Outside the United States, the risks of DC were significantly different in GAD.
and MDD studies.

Table 14: TKC Test 3 Partitions for Risks of DC with Study Location

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Weight Function</th>
<th>Non-US</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test Statistic</td>
<td>P-value</td>
<td>Test Statistic</td>
</tr>
<tr>
<td>GAD</td>
<td>$W_{jk}^0$</td>
<td>29.72</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^1(\rho = 1)$</td>
<td>29.84</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^2(\rho = 1, r = 0)$</td>
<td>30.75</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^3(\rho = 1)$</td>
<td>31.21</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>MDD</td>
<td>$W_{jk}^0$</td>
<td>11.51</td>
<td>0.0032</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^1(\rho = 1)$</td>
<td>11.38</td>
<td>0.0034</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^2(\rho = 1, r = 0)$</td>
<td>9.18</td>
<td>0.0101</td>
</tr>
<tr>
<td></td>
<td>$W_{jk}^3(\rho = 1)$</td>
<td>13.52</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

5.4 Regression Modeling Results

To determine significant risk factors in both the Cox CSH and Fine & Gray regression models, a general model selection strategy utilizing differences between
Figure 11: CCSH Curves for MDD and United States

model likelihoods outlined by Collett (Collett, 2003) will be used. Let $\hat{L}$ denote the estimated maximum likelihood function for the model. First, models that contain variables of interest one at a time are fitted, and the value of $-2 \log \hat{L}$ for the model including the variable is compared with the value of $-2 \log \hat{L}$ for the model not including any variables. The variables that significantly reduce the value of this statistic are then fitted together. To determine the variables that do not significantly contributed to this full model, the difference in $-2 \log \hat{L}$ between the full model and the model omitting a particular variable is computed, and variables that lead to a significant increase in $-2 \log \hat{L}$ when omitted are retained in the model. The variables that were eliminated in the first step are added to the model from the second step one at a time, and the variables that significantly decrease the difference in $-2 \log \hat{L}$ are retained in the full model. Lastly, the variables
included in the full model are tested to make sure that they cannot be omitted in the model without significantly increasing the value of $-2 \log \hat{L}$ and that no variable not included significantly decreases the value of $-2 \log \hat{L}$.

The following risk factors will be considered in the regression models: treatment assignment (placebo or treatment group), study location (United States and outside United States), race (Caucasian and non-Caucasian), gender, age at start of treatment, study duration, standardized early percent change in efficacy, and standardized early maximum event severity. The percent change in efficacy at visit week $i$ is calculated as

$$PCTCHANGE(i) = 100 \times \frac{EFFVAL(i) - BASEVAL}{BASEVAL} \quad (5.1)$$

where $PCTCHANGE(i)$ is the percent change in efficacy at visit week $i$, $EFFVAL(i)$ is the efficacy value at visit week $i$, and $BASEVAL$ is the baseline visit efficacy value. The maximum event severity at visit week $i$ is calculated as

$$MAX\_SEV(i) = \text{MAX}\{EVENT\_SEV(1), ..., EVENT\_SEV(p)\} \quad (5.2)$$

where $EVENT\_SEV(j)$ is the $j$th adverse event severity rating at visit week $i$.

Tables 15 and 16 contain the final risk factors selected for the risk of DC due to AE in the Fine & Gray and Cox CSH models, respectively. In both regression models, an increase of one standard deviation in standardized maximum event
severity increases the risk of DC due to AE by approximately 110% in GAD studies and 80% in MDD studies. Patients assigned to the treatment group had an increased risk of DC due to AE by at least 50% in all disease states.

Table 15: CRR Model Selection Results for DC due to AE

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Risk Factor</th>
<th>Subdistribution HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>Standardized PCTCHANGE</td>
<td>1.19</td>
<td>1.06 1.33</td>
</tr>
<tr>
<td></td>
<td>Standardized MXSEVP</td>
<td>2.12</td>
<td>1.80 2.50</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.68</td>
<td>0.49 0.94</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>2.28</td>
<td>1.57 3.30</td>
</tr>
<tr>
<td>MDD</td>
<td>Age (yrs)</td>
<td>1.02</td>
<td>1.01 1.03</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.12</td>
<td>1.00 1.24</td>
</tr>
<tr>
<td></td>
<td>Standardized MXSEVP</td>
<td>1.79</td>
<td>1.58 2.03</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>1.68</td>
<td>1.24 2.29</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.99</td>
<td>0.70 1.40</td>
</tr>
</tbody>
</table>

Table 16: Cox PH Model Selection Results for DC due to AE

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Risk Factor</th>
<th>Hazard Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>Standardized PCTCHANGE</td>
<td>1.20</td>
<td>1.08 1.35</td>
</tr>
<tr>
<td></td>
<td>Standardized MXSEVP</td>
<td>2.19</td>
<td>1.85 2.59</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.65</td>
<td>0.47 0.90</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>2.26</td>
<td>1.55 3.29</td>
</tr>
<tr>
<td>MDD</td>
<td>Age (yrs)</td>
<td>1.02</td>
<td>1.01 1.03</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.13</td>
<td>1.01 1.27</td>
</tr>
<tr>
<td></td>
<td>Standardized MXSEVP</td>
<td>1.82</td>
<td>1.61 2.06</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>1.66</td>
<td>1.22 2.25</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.93</td>
<td>0.65 1.32</td>
</tr>
</tbody>
</table>

Tables 17 and 18 contain the final risk factors selected for the risk of DC due to LOE in the Fine & Gray and Cox CSH models, respectively. In both regression models, an increase of one standard deviation in standardized percent change in efficacy increases the risk of DC due to LOE by approximately 35% in
GAD studies and 24% in MDD studies. Patients assigned to the placebo group had an increased risk of DC due to LOE by at least 50% in all disease states. Patients treated in the United States were at lower risk of DC due to LOE in GAD and MDD studies than patients treated outside the United States in these disease states.

Table 17: CRR Model Selection Results for DC due to LOE

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Risk Factor</th>
<th>Subdistribution HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>Age (yrs)</td>
<td>1.02</td>
<td>1.00 1.03</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.35</td>
<td>1.19 1.53</td>
</tr>
<tr>
<td></td>
<td>Country-United States</td>
<td>0.35</td>
<td>0.21 0.59</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>0.20</td>
<td>0.12 0.33</td>
</tr>
<tr>
<td>MDD</td>
<td>Standardized PCTCHANGE</td>
<td>1.24</td>
<td>1.14 1.34</td>
</tr>
<tr>
<td></td>
<td>Standardized MXSEVP</td>
<td>0.90</td>
<td>0.80 1.02</td>
</tr>
<tr>
<td></td>
<td>Study Duration (weeks)</td>
<td>1.13</td>
<td>1.01 1.26</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>1.11</td>
<td>0.77 1.60</td>
</tr>
<tr>
<td></td>
<td>Country-United States</td>
<td>0.42</td>
<td>0.33 0.54</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>0.46</td>
<td>0.36 0.58</td>
</tr>
</tbody>
</table>

Table 18: Cox PH Model Selection Results for DC due to LOE

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Risk Factor</th>
<th>Subdistribution HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>Age (yrs)</td>
<td>1.02</td>
<td>1.00 1.03</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.37</td>
<td>1.18 1.59</td>
</tr>
<tr>
<td></td>
<td>Country-United States</td>
<td>0.34</td>
<td>0.21 0.58</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>0.21</td>
<td>0.13 0.34</td>
</tr>
<tr>
<td>MDD</td>
<td>Standardized PCTCHANGE</td>
<td>1.28</td>
<td>1.17 1.40</td>
</tr>
<tr>
<td></td>
<td>Country-United States</td>
<td>0.41</td>
<td>0.32 0.53</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>0.45</td>
<td>0.36 0.57</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>1.05</td>
<td>0.72 1.51</td>
</tr>
</tbody>
</table>

Tables 19 and 20 contain the final risk factors selected for the risk of DC
due to other reasons in the Fine & Gray and Cox CSH models, respectively. Early standardized percent change in efficacy was a marginally significant predictor of DC due to other reasons in DPNP studies. Caucasian patients had a decreased risk of DC due to other reasons by approximately 30% and 40% in GAD and MDD studies, respectively.

Table 19: CRR Model Selection Results for DC due to OTH

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Risk Factor</th>
<th>Subdistribution HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>Age (yrs)</td>
<td>0.97</td>
<td>0.97 0.98</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.71</td>
<td>0.56 0.90</td>
</tr>
<tr>
<td></td>
<td>Country-United States</td>
<td>1.85</td>
<td>1.48 2.31</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.06</td>
<td>0.97 1.18</td>
</tr>
<tr>
<td>MDD</td>
<td>Age (yrs)</td>
<td>0.97</td>
<td>0.97 0.98</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.59</td>
<td>0.48 0.74</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.05</td>
<td>0.96 1.15</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>1.28</td>
<td>1.05 1.57</td>
</tr>
</tbody>
</table>

Table 20: Cox PH Model Selection Results for DC due to OTH

<table>
<thead>
<tr>
<th>Disease State</th>
<th>Risk Factor</th>
<th>Subdistribution HR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAD</td>
<td>Age (yrs)</td>
<td>0.97</td>
<td>0.97 0.98</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.69</td>
<td>0.54 0.88</td>
</tr>
<tr>
<td></td>
<td>Country-United States</td>
<td>1.83</td>
<td>1.45 2.31</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.10</td>
<td>0.99 1.22</td>
</tr>
<tr>
<td>MDD</td>
<td>Age (yrs)</td>
<td>0.97</td>
<td>0.97 0.98</td>
</tr>
<tr>
<td></td>
<td>Treatment Duration (weeks)</td>
<td>0.90</td>
<td>0.82 0.98</td>
</tr>
<tr>
<td></td>
<td>Race-Caucasian</td>
<td>0.59</td>
<td>0.48 0.74</td>
</tr>
<tr>
<td></td>
<td>Group-Treatment Group</td>
<td>1.26</td>
<td>1.03 1.55</td>
</tr>
<tr>
<td></td>
<td>Standardized PCTCHANGE</td>
<td>1.08</td>
<td>0.99 1.17</td>
</tr>
</tbody>
</table>

In these regression models for categories of discontinuation, the hazard ratio estimates of the risk factors for the Fine & Gray and Cox CSH regression
models were in close agreement with each other.
CHAPTER VI

SUMMARY AND CONCLUSIONS

6.1 Summary

The focus of this dissertation has been conducting statistical inference on different competing risks applied to patient discontinuation in clinical trials. The competing risks methodology goes beyond the traditional analysis of patient discontinuation by analyzing risk factors for each category of discontinuation and determining differences between the risks of discontinuation without introducing bias in the tests and model procedures.

The competing risks methodology analyzes the joint distribution of the time to discontinuation and the type of discontinuation. The two fundamental parameters of interest, the cause-specific hazard and the cumulative incidence function, are directly estimable from the observed data. Plots of the estimated cumulative cause-specific hazard or cumulative incidence functions over the duration of the experiment help to visualize differences between the competing risks and the form of the hazards for these competing risks. Inference procedures like the KLY test and the three TKC tests provide ways to directly compare the competing risks of discontinuation by comparing cause-specific hazards, and in the case of the TKC tests, incorporate grouping factors into the tests. Both proce-
dures incorporate weight functions that are appropriate in different conditions. The Fine & Gray regression model for the hazard of subdistribution provides a direct analog to the Cox proportional hazards model for modelling the cumulative incidence function for each competing risks of discontinuation. Counting process and martingale theory provide theoretical justification of the inference and modelling procedures as a special case of event history analysis.

The TKC test procedures for comparing the risks of discontinuation have distinct advantages over the KLY testing procedure. First, the TKC procedure is able to incorporate more than two competing risks, and it is also able to determine differences between groups. The KLY procedure compares two competing risks and does not take group factors into account. The construction of the TKC test statistics follows a unified framework using martingale theory with convergence to the well-known chi-square distribution, while the two-tailed version of the KLY test statistic converges to the range of a standard Brownian motion, with the probability density function described by Feller (Feller, 1951). Lastly, partitions of the TKC test statistics for the procedures that compare different group levels within a competing risk and compare competing risks within a group level can be examined to see which particular grouping levels or competing risks contain significant differences, similar to comparing means within group levels in an analysis of variance framework.
6.2 Conclusions

The results from the simulation study of the different weight functions used in the TKC tests show that a type I error rate close to the 5% level is maintained using the first and third weight functions with \( p = 1 \), and the second weight function with \( p = 1 \) and \( r = 0 \). In light and moderate censoring conditions, the second weight function with \( p = 1 \) and \( r = 0 \) and the third weight function with \( p = 1 \) provided slight gains in power compared to the default weight function.

The TKC test procedures found significant differences between the risks of discontinuation due to adverse events and lack of efficacy between and within treatment groups. Plots of the cumulative cause specific hazard and estimates of the hazard ratio from the Fine & Gray and Cox cause specific hazard regression models provided evidence that patients in the placebo group were at higher risk of discontinuation due to lack of efficacy and patients in the treatment group were at higher risk of discontinuation due to adverse events in both disease states. In almost every case, each of the weight functions lead to the same conclusion in the TKC tests.

6.3 Additional Research

There are several research topics that have the potential to enhance the competing risks analysis of patient discontinuation data. While the competing risks methodology is able to account for possible dependence between the com-
peting risks, currently there is no unified procedure to quantify the dependence between different competing risks. There are procedures available to test the dependence between the time to event and type of event observed (Dewan et al., 2004) and (Gasbarra et al., 2006), however these procedures are not readily available in standard software packages. Future research could determine the appropriate conditions for the different weight functions that were utilized in the TKC procedures of this dissertation. For example, which of the weight functions are appropriate for early differences between the competing risks, late differences between the competing risks, or the situation when the cumulative CSH curves cross at one or more times in the study. Lastly, the current version of the Fine & Gray model is not able to handle continuous time-dependent covariates in the model. Currently research using the multi-state modelling approach is being conducted to form regression models for competing risks that can have continuous time-dependent covariates (Beyersmann & Schumacher, 2008) and to develop non-parametric inference procedures that test the effect of a continuous covariate on a competing risks (Peng & Fine, 2008).
REFERENCES


Heckman, J. (1976). The Common Structure of Statistical Models of Truncation,


APPENDIX A

PROGRAM CODE

A.1: Simulation Driver Program

```r
### Simulation program written 3--2--09
#
#
## Using 3 risks and 3 groups
##
### Set weight function using weight= assignment
### Set distribution for random generation using dist= assignment ("mvn", "exp", "weibull")
### Set number of simulations using n_sim= assignment
### Set sample size of groups using n= assignment

cat("simulation_log\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="", append=FALSE)
source("R_funcs/lognsim.R")
source("R_funcs/simtest.R")

weight=l; dist="weibull"; n_sim=1000; n=200

thetamatrix1 = matrix(c(1,1,1,1,1,1,1,1,1), byrow=T, nrow=3)
thetamatrix2 = matrix(c(0.5,0.5,0.5,1,1,2,2,2,2), byrow=T, nrow=3)
thetamatrix3 = matrix(c(0.5,1,2,0.5,1,2,0.5,1,2), byrow=T, nrow=3)
J = nrow(thetamatrix1)

#thetavec = seq(0.7,1.5,by=0.1)
thetavec=c(0.5,1,2,4,8)
theticens=c(1)

#seedstart=1
index=1
indexrho=0
end=10
```
if (weight==1) {
    rhovec_weight1 = c(0,1)
    #seedarray_weight1 = array(NA, dim=c(n_sim, J, (length(rhovec_weight1)*length(thetavec)))))
    resultarray_thetamatrix1_weight1 <- resultarray_thetamatrix2_weight1 <- resultarray_thetamatrix3_weight1 <- array(NA, dim=c(length(rhovec_weight1), length(thetavec), 3),
        dimnames=list(rhovec_weight1, thetavec, c("Test1", "Test2", "Test3")))
    avgcensorarray_thetamatrix1_weight1 <- avgcensorarray_thetamatrix2_weight1 <- avgcensorarray_thetamatrix3_weight1 <- matrix(NA, nrow=length(rhovec_weight1), ncol=length(thetavec),
        dimnames=list(rhovec_weight1, thetavec))
    for (rho in rhovec_weight1) {
        indextheta = 1
        indexrho = indexrho+1
        for (theta in thetavec) {
            #seedarray_weight1[,index] = matrix(c(seq(seedstart, (seedstart + (J*n_sim)-1))),
                nrow=n_sim, ncol=J, byrow=T)
            #seedarray_weight1[,index] = matrix(randomSequence(seedstart, seedstart + (J*n_sim)-1),
                nrow=n_sim, ncol=J, byrow=T)
            cat("rho_", rho, ", theta_", theta, ", _sim_for_thetamatrix1\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="", append=TRUE)
            simresult_thetamatrix1 = simtest(dist=dist, theta=thetamatrix1, n=n, thetacens=thetacens, rho=rho, r=0, weight=1, n_sim=n_sim)
            cat("rho_", rho, ", theta_", theta, ", _sim_for_thetamatrix2\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="", append=TRUE)
            simresult_thetamatrix2 = simtest(dist=dist, theta=thetamatrix2, n=n, thetacens=thetacens, rho=rho, r=0, weight=1, n_sim=n_sim)
            cat("rho_", rho, ", theta_", theta, ", _sim_for_thetamatrix3\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="", append=TRUE)
            simresult_thetamatrix3 = simtest(dist=dist, theta=thetamatrix3, n=n, thetacens=thetacens, rho=rho, r=0, weight=1, n_sim=n_sim)
            avgcensorarray_thetamatrix1_weight1[indexrho,indextheta] = simresult_thetamatrix1$avg_censored
            avgcensorarray_thetamatrix2_weight1[indexrho,indextheta] = simresult_thetamatrix2$avg_censored
            avgcensorarray_thetamatrix3_weight1[indexrho,indextheta] = simresult_thetamatrix3$avg_censored
            resultarray_thetamatrix1_weight1[indexrho,indextheta,1] = simresult_thetamatrix1$prop_typeone_test1
            resultarray_thetamatrix1_weight1[indexrho,indextheta,2] = simresult_thetamatrix1$prop_typeone_test2
            resultarray_thetamatrix1_weight1[indexrho,indextheta,3] = simresult_thetamatrix1$prop_typeone_test3
resultarray.thetamatrix2.weight1[indexrho,indextheta,1] = simresult.thetamatrix2$prop.typeone_test1
resultarray.thetamatrix2.weight1[indexrho,indextheta,2] = simresult.thetamatrix2$prop.typeone_test2
resultarray.thetamatrix2.weight1[indexrho,indextheta,3] = simresult.thetamatrix2$prop.typeone_test3
resultarray.thetamatrix3.weight1[indexrho,indextheta,1] = simresult.thetamatrix3$prop.typeone_test1
resultarray.thetamatrix3.weight1[indexrho,indextheta,2] = simresult.thetamatrix3$prop.typeone_test2
resultarray.thetamatrix3.weight1[indexrho,indextheta,3] = simresult.thetamatrix3$prop.typeone_test3

#seedstart = seedstart + (J*n-sim)
index = index + 1
indextheta = indextheta + 1
}
}
#seedlast_weight1 = seedstart - 1
}

if (weight==2) {
    rhovec.weight2 = c(0,1)
    #seedarray.weight2 = array(NA, dim=c(n_sim, J, (length(rhovec.weight2)*length(thetavec))))
    resultarray.thetamatrix1.weight2 <= resultarray.thetamatrix2.weight2 <= resultarray.thetamatrix3.weight2 <= array(NA, dim=c(length(rhovec.weight2),length(thetavec), 3),
dimnames=list(rhovec.weight2, thetavec, c("Test1", "Test2", "Test3"))
    avgcensorarray.thetamatrix1.weight2 <= avgcensorarray.thetamatrix2.weight2 <= avgcensorarray.thetamatrix3.weight2 <= matrix(NA, nrow=length(rhovec.weight2), ncol=length(thetavec),
dimnames=list(rhovec.weight2, thetavec))
    for (rho in rhovec.weight2) {
        indextheta = 1
        indexrho = indexrho+1
        for (theta in thetavec) {
            #seedarray.weight2[,index] = matrix(c(seq(seedstart,(seedstart + (J*n_sim)-1))),
nrow=n_sim, ncol=J, byrow=T)
            #seedarray.weight2[,index] = matrix(randomSequence(seedstart,seedstart + (J*n_sim)
-1), nrow=n_sim, ncol=J, byrow=T)
            cat("rho", rho, "_theta", theta, "_sim", for.thetamatrix1\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="\n", append=TRUE)
simresult_thetamatrix1 = simtest(dist=dist, theta=thetamatrix1, n=n, thetacens = thetacens, rho=rho, r=(1-rho), weight=2, n_sim=n_sim)

cat("rho_{\cdot}, "), rho, "_theta_{\cdot}, "theta, "_\cdot_sim_for_thetamatrix2\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="", append=TRUE)

simresult_thetamatrix2 = simtest(dist=dist, theta=thetamatrix2, n=n, thetacens = thetacens, rho=rho, r=(1-rho), weight=2, n_sim=n_sim)

cat("rho_{\cdot}, "), rho, "_theta_{\cdot}, "theta, "_\cdot_sim_for_thetamatrix3\n", file="/home/eric/dissertation_backup/rsim/siminfo.txt", sep="", append=TRUE)

simresult_thetamatrix3 = simtest(dist=dist, theta=thetamatrix3, n=n, thetacens = thetacens, rho=rho, r=(1-rho), weight=2, n_sim=n_sim)

avg_censorarray_thetamatrix1_weight2 [indexrho, indextheta] = simresult_thetamatrix1$avg_censored

avg_censorarray_thetamatrix2_weight2 [indexrho, indextheta] = simresult_thetamatrix2$avg_censored

avg_censorarray_thetamatrix3_weight2 [indexrho, indextheta] = simresult_thetamatrix3$avg_censored

resultarray_thetamatrix1_weight2 [indexrho, indextheta, 1] = simresult_thetamatrix1$prop_typeone_test1

resultarray_thetamatrix1_weight2 [indexrho, indextheta, 2] = simresult_thetamatrix1$prop_typeone_test2

resultarray_thetamatrix1_weight2 [indexrho, indextheta, 3] = simresult_thetamatrix1$prop_typeone_test3

resultarray_thetamatrix2_weight2 [indexrho, indextheta, 1] = simresult_thetamatrix2$prop_typeone_test1

resultarray_thetamatrix2_weight2 [indexrho, indextheta, 2] = simresult_thetamatrix2$prop_typeone_test2

resultarray_thetamatrix2_weight2 [indexrho, indextheta, 3] = simresult_thetamatrix2$prop_typeone_test3

resultarray_thetamatrix3_weight2 [indexrho, indextheta, 1] = simresult_thetamatrix3$prop_typeone_test1

resultarray_thetamatrix3_weight2 [indexrho, indextheta, 2] = simresult_thetamatrix3$prop_typeone_test2

resultarray_thetamatrix3_weight2 [indexrho, indextheta, 3] = simresult_thetamatrix3$prop_typeone_test3

#seedstart = seedstart + (J*n_sim)
index = index + 1
indextheta = indextheta + 1

```
#seclast.weight3 = seedstart - 1

if (weight==3) {
  rhovec.weight3 = c(0,0.5,1)
  #seed.array.weight3 = array(NA, dim=c(n.sim, J, length(rhovec.weight3)*length(thetavec)))
  resultarray.thetamatrix1.weight3 <- resultarray.thetamatrix2.weight3 <- resultarray.thetamatrix3.weight3 <- array(NA, dim=c(length(rhovec.weight3), J, length(thetavec)), 
             dimnames=list(rhovec.weight3, thetavec, c("Test1", "Test2", "Test3")))
  avgcensorarray.thetamatrix1.weight3 <- avgcensorarray.thetamatrix2.weight3 <- avgcensorarray.thetamatrix3.weight3 <- matrix(NA, nrow=length(rhovec.weight3), ncol=length(thetavec), 
             dimnames=list(rhovec.weight3, thetavec))
  for (rho in rhovec.weight3) {
    indexrho = indexrho+1
    for (theta in thetavec) {
      #seed.array[,, index] = matrix(seq(seedstart, (seedstart + (J*n.sim)-1)),
                                     nrow=n.sim, ncol=J, byrow=T)
      #seed.array[,, index] = matrix(randomSequence(seedstart, seedstart + (J*n.sim)-1), nrow=n.sim, ncol=J, byrow=T)
      cat("rho", rho, "-theta ", theta, "sim for thetamatrix1
" , file="/home/eric/dissertation_backup/rsim/siminfo.txt" , sep=" ", append=TRUE)
      simresult.thetamatrix1 = simtest(dist=dist, theta=thetamatrix1, n=n, thetacens=thetacens, rho=rho, r=0, weight=3, n_sim=n.sim)
      cat("rho", rho, "-theta ", theta, "sim for thetamatrix2
" , file="/home/eric/dissertation_backup/rsim/siminfo.txt" , sep=" ", append=TRUE)
      simresult.thetamatrix2 = simtest(dist=dist, theta=thetamatrix2, n=n, thetacens=thetacens, rho=rho, r=0, weight=3, n_sim=n.sim)
      cat("rho", rho, "-theta ", theta, "sim for thetamatrix3
" , file="/home/eric/dissertation_backup/rsim/siminfo.txt" , sep=" ", append=TRUE)
      simresult.thetamatrix3 = simtest(dist=dist, theta=thetamatrix3, n=n, thetacens=thetacens, rho=rho, r=0, weight=3, n_sim=n.sim)
      avgcensorarray.thetamatrix1.weight3[indexrho, indextheta] = simresult.thetamatrix1$avg_censored
      avgcensorarray.thetamatrix2.weight3[indexrho, indextheta] = simresult.thetamatrix2$avg_censored
      avgcensorarray.thetamatrix3.weight3[indexrho, indextheta] = simresult.thetamatrix3$avg_censored
      resultarray.thetamatrix1.weight3[indexrho, indextheta,1] = simresult.thetamatrix1$prop_typeone.test
    }
  }
}
resultarray\_theta\_matrix\_weight3[index\_rho,\ index\_theta, 2] = simresult\_theta\_matrix\_weight3

prop\_typeone\_test2
resultarray\_theta\_matrix\_weight3[index\_rho,\ index\_theta, 3] = simresult\_theta\_matrix\_weight3

prop\_typeone\_test3
resultarray\_theta\_matrix\_weight3[index\_rho, \ index\_theta, 2] = simresult\_theta\_matrix\_weight3

prop\_typeone\_test2
resultarray\_theta\_matrix\_weight3[index\_rho, \ index\_theta, 3] = simresult\_theta\_matrix\_weight3

prop\_typeone\_test3

\#seedstart = seedstart + (J*n-sim)
index = index + 1
indextheta = indextheta + 1

\}
\}
\#seedlast\_weight3 = seedstart - 1

save.image(file=paste("/home/eric/dissertation-backup/rsim/simulation\_results\_nsim" ,n-sim,"_dist\_" ,dist,"_groupsizes\_" ,n,"_censrate\_" ,thetacens,"_weightfxn\_" ,weight,"_space\_RData", sep=""))

A.2: Data Simulator

lognsim <- function(dist, theta, n, thetacens) {
  K = ncol(theta)
  J = nrow(theta)
  data = matrix(NA, nrow = n, ncol = J)
  alldata = array(dim = c(n, 3, K))
  simdata <- simcause <- rep(NA, n)
  cens = rexp(n, rate = thetacens)
  for (k in 1:K) {
    group = c(rep(k, n))
    if (dist == "exp") {
      for (j in 1:J) {
        data[, j] = rexp(n, rate = theta[j, k])
      }
    }
  }
  return(data)
}
else if (dist="mvn") {
    library("compositions")
    set.seed(seeds[l])
    data = rnorm.rplus(n,theta[k],var)
}
else if (dist="weibull") {
    for (j in 1:J) {
        data[,j] = rweibull(n, shape=theta[j,k], scale=1)
    }
}

min = apply(data,1,min)
cause = apply(data,1,which.min)
for (i in 1:n) {
    simdata[i] = min(min[i], cens[i])
    simcause[i] = cause[i] * (min[i] < cens[i])
}
alldata[,k] = cbind(simdata, simcause, group)
}
alldatanew <- NULL
alldatanew = as.data.frame(alldatanew)
for (k in 1:K) {
    alldatanew = rbind(alldatanew, alldatanew[,k])
}
colnames(alldatanew) = c("time", "cause", "group")
timeslist = split(alldatanew$time, alldatanew$group)
causelist = split(alldatanew$cause, alldatanew$group)
list(alldatanew = alldatanew, timeslist=timeslist, causelist=causelist, causeunlist=
     alldatanew$cause)
}

A.3: Test Simulator

simtest <- function(dist, theta, n, thetacens, rho, r, weight, n_sim) {
    # took out seedmatrix parameter

    #source("R_funcs/lognsim.R")
    source("R_funcs/FNweight_weights.R")
    iteration <- pctcens <- teststat1 <- teststat2 <- teststat3 <- pvt.test1 <- pvt.test2 <- pvt.test3 <- power1 <- power2 <- power3 <- df <- ng <- quantile <- rep(NA, n_sim)
    for (i in 1:n_sim) {
             sep="",append=TRUE)
A.4: KLY Test Procedure Program

kly_weightfxn=function(time,cens,weight)
{
  ## based on Kochar, Lam and Yip #
  ## Lifetime Data Analysis, 2002 #
  
  #sim = lognsim(dist=dist,theta=theta,n=n,thetak=n,thetacens=thetak,seed=
  #seedmatrix[i,])
sim = lognsim(dist=dist,theta=theta,n=n,thetak=n,thetacens=thetak)
K = length(sim	imes list)
J = max(unlist(sim	imes causelist))
results = weight.test(X=sim	imes list,d=sim	imes causelist,rho=rho,r=r,weight=
  weight)
#ccsh = resultsSccsh
#causelist = unsplit(sim	imes causelist)
pctcens[i] = round(length(sim	imes causeunlist[1]causeunlist==0))/length(sim
  causeunlist),3)
teststat1[i] = results@teststats[1]
teststat2[i] = results@teststats[2]
teststat3[i] = results@teststats[3]
 pv_test1[i] = results$pvalues[1]
pv_test2[i] = results$pvalues[2]
 pv_test3[i] = results$pvalues[3]
power1[i] = pv_test1[i] <= 0.05
power2[i] = pv_test2[i] <= 0.05
power3[i] = pv_test3[i] <= 0.05
  # results$pvalues[1] for test 1, results$pvalues[2] for test2, results$pvalues
  [3] for test 3
}
  
  avg.censored = mean(pctcens)
  prop_typeone_test1 = round(sum(power1)/n_sim,8)
  prop_typeone_test2 = round(sum(power2)/n_sim,8)
  prop_typeone_test3 = round(sum(power3)/n_sim,8)
  #prop_typeone_test1 = round(mean(pv_test1 < 0.05),6)
  #prop_typeone_test2 = round(mean(pv_test2 < 0.05),6)
  #prop_typeone_test3 = round(mean(pv_test3 < 0.05),6)
  list(pctcens=pctcens, avg.censored=avg.censored, teststat1=teststat1, teststat2=teststat2,
   teststat3=teststat3, pv_test1=pv_test1, pv_test2=pv_test2, pv_test3=pv_test3, prop
   typeone_test1=prop_typeone_test1, prop_typeone_test2=prop_typeone_test2, prop_typeone
   test3=prop_typeone_test3)
114

# time=time to first event  
# cens=censor variable  
# l=one type of event  
# 2=another type of event  
# 0=the rest of events and#  
# censored records  

dd0=table(time,cens)

if (sum(cens==0)==0) {dd=cbind(rep(0,nrow(dd0)),dd0)}
else {dd=dd0}

if (sum(cens==l)==0 | sum(cens==2)==0 | sum(cens>2)>0)
{stop("Codes-for-events should be 1 or 2")}

n=length(time)

nt=apply(dd,1,sum)

ntt=cumsum(nt)

#atrisk=n—ntt

# wl(u) = Ybar(u)
if (weight==1) {

  lnti=(dd[,3]-dd[,2])
  sigma=sqrt(n*sum(cens>0))
}

# w2(u) = Ybar'2(u)
else if (weight==2) {

  lnti=atrisk*(dd[,3]-dd[,2])
  sigma=sqrt(n*sum(sapply(atrisk, function(x) x^2)*(dd[,3]+dd[,2])))
  #sigma=sqrt(n)*sum(atrisk*(dd[,3]+dd[,2]))
}

# w3(u) = Ybar(u)*Nbar(u—)
else if (weight==3) {

  ddi=(dd[,3]+dd[,2])
  ddsum=cumsum(ddi)
\[ \text{inti} = \text{ddsum} \ast (\text{dd}[3] - \text{dd}[2]) \]
\[ \text{sigma} = \sqrt{n \ast \text{sum(sapply(ddsum, function(x) x^2) \ast (dd[3] + dd[2]))}} \]

```r
# w4(u) = Ybar'2(u) \ast Nbar(u-) else if (weight==4) {
  ddi = (dd[3] + dd[2])
  ddsum = cumsum(ddi)
  inti = atrisk \ast \text{ddsum} \ast (\text{dd}[3] - \text{dd}[2])
  sigma = \sqrt{n \ast \text{sum(sapply(atrisk, function(x) x^2) \ast sapply(ddsum, function(x) x^2) \ast (dd[3] + dd[2]))}}
}

int = cumsum(inti)
k = nrow(dd)
maxint = vector("numeric", k)

# \text{sigma} = \sqrt{n \ast \text{sum(cens > 0)}}

for (i in 1:k)
{
  maxint[i] = \text{max(abs(int - int[i])})
}

cnw = sqrt(n) \ast \text{max(maxint)}/\text{sigma}

## for C, (without absolute value)
penw = function(cc)
{
  k = c(0:15)
  parti = (-1)^k \ast \exp(-pi/2 \ast (2\ast k+1)^2 / (8 \ast cc^2)) / (2\ast k+1)
  p = sum(parti) \ast 4/pi
  return(p)
}

## for Cstar (with absolute value)
ph = function(s)
{
  f1 = function(x1, x2, k, xi)
  {
    k = c(1:k)
    x = c((x1\ast xi) : (x2\ast xi))/xi
  }
  return(f1)
}
```

```environment```
n = \text{length}(x)

h = \text{vector}("\text{numeric}", n)

\text{for} \ (i \ \text{in} \ 1:n) \ 
\{ 
\quad h[i] = \text{sum}(8 \times (-1)^i \times (i - 1) \times k_i^2 \times \text{dnorm}(k_i \times x[i])) 
\}

\text{parth} = \text{sum}(h) / xi 
\text{return}(\text{parth}) 

\text{if} \ (z \leq 0.1) \ 
\quad \text{ph} = f(0.01, z, 1000, 100) 
\text{else if} \ (z \leq 0.5) \ 
\quad \text{ph} = f(0.01, 0.1, 1000, 100) + f(0.11, z, 100, 100) 
\text{else ph} = f(0.01, 0.1, 1000, 100) + f(0.11, 0.5, 100, 100) + f(0.51, z, 30, 1000) 
\text{return}(\text{ph}) 
\}

\text{pvalue} = \text{round}(1 - \text{ph}(\text{cnw}), 4) 
\# \text{if} \ (\text{pvalue} \geq 0.0001) \ \text{r} = \text{data.frame}(\text{cnw}, \text{pvalue}) 
\# \text{if} \ (\text{pvalue} < 0.00001) \ \text{r} = \text{data.frame}(\text{cnw}, \text{pvalue}="p-value < 0.00001") 
\# \text{row.names}(r)="" 
\# \text{return}(r) 
\text{if} \ (\text{pvalue} \geq 0.0001) \ \text{list}(\text{cnw} = \text{cnw}, \text{pvalue} = \text{pvalue}) 
\text{else if} \ (\text{pvalue} < 0.0001) \ \text{list}(\text{cnw} = \text{cnw}, \text{pvalue} = "<0.0001") 
\}

A.5: TKC Test Procedure Program

# File name : FNweight_weights.R 
# Author : Eric Nantz 
# Version : 1.0, January 20, 2009 
# based on R Version 2.8.1 () 
# Required function: ginv from MASS (http://www.stats.oz.ac.uk/pub/MASS4/) 
# 
# Function : weight.test 
# 
# <input> : X = list of lifetime (group, 1,2...K)
# d = list of censor (0) and type of cause (1, 2, ..., m)
# m = number of competing risks
# rho = parameter used in weight functions 1, 2, 3 (default = 0)
# r = additional parameter used in weight function 2 (default = 0)
# weight = type of weight function to use in test statistic (default = 1)
# 1: Wjk = Yk(t)\* [\hat{Y}/n]^\rho
# 2: Wjk = Yk(t) \* (\hat{S}(t))^{\rho} \* (1 - \hat{S}(t))^{\rho}
# 3: Wjk = Yk(t) \* (1 - \hat{S}(t))^{\rho}
# <output> : T1 : test statistics under H01
# T2 : test statistics under H02 (if k>1)
# T3 : test statistics under H03
# ccsh : array
# pvalueT1 : pvalue for test statistic under H01
# pvalueT2 : pvalue for test statistic under H02 (if k>1)
# pvalueT3 : pvalue for test statistic under H03

## Using rat data from Hoel
## X = list ( rats ~ allgroups & time { rats = allgroups & time [ rats ~ allgroups & group == 1], rats ~ allgroups & time [ rats ~ allgroups & group == 2])
## d = list ( rats ~ allgroups & code [ rats ~ allgroups & group == 1], rats ~ allgroups & code [ rats ~ allgroups & group == 2])
## m leave as undefined (the fn takes care of figuring it out)
## rho and t leave as defaults of rho=0 and t=Inf (NOTE: Inf is positive infinity)

weight.test <- function(X, d, m, rho=0, r=0, weight=1, t=Inf , type="ccsh", plot="F", legend="T", datatext=0, vartext=0, dscatplot=0, groupplot=0) {
  library("MASS")
  TINY <- sqrt(.Machine$double.eps)
  K <- length(X)
  n.k <- c(rep(0, K))
  for(k in 1:K) { n.k[k] = length(X[[k]]) }
  if (missing(m)) m <- max( unlist(d) )
  x <- unlist(X)
  n <- length(x)

  ## s <- x[x < t+TINY] ### t
  ## this creates a vector s with all of the unique failure times (i.e. no duplicate times)
  (173 unique times in rat data)
  s <- unique ( x[x < t+TINY] ) ### t
  #s <- unique ( x[(out - TINY < x & x < t+TINY] ) ### t
## an extra 0 is added to the beginning of the s vector for using the diff function in the
test process estimation later on

```r
if (sort(s)[1] == 0) {
  s <- sort(s)
} else {
  s <- c(0, sort(s))
}
ls <- length(s)
```

### N is a multidimensional array for number of failures composed of Is matrices, each with
dimension m×k
### Each submatrix corresponds to a sorted failure time, with the columns for each group and
rows for the number of failures of each type up to that sorted time
### Example: term [5,1,174] is the 174th matrix's 2nd row, 1st column element

```r
N <- array( dim=c(m,K,ls) )
```

### Y is a 2-dimensional array for number of patients at risk with K rows (each row
  corresponds to group 1,...,k)
### And (ls-1) columns correspond to the sorted failure times

```r
Y <- array( dim=c(K,ls-1) )
```

### this series of loops

```r
for ( k in 1:K ) {
  x <- X[[k]]
  # sum(x > s[si]-TINY) calculates how many of the kth group's failure times are greater
  # than the si-th sorted failure time (hence number at risk)
  for (si in 2:ls) Y[k,si-1] <- sum( x > s[si]-TINY )
  for ( j in 1:m ) {
    for (si in 1:ls) N[j,k,si] <- sum( (x<s[si]+TINY)*(d[[k]]==j) )
  }
}
```

### Y. is matrix for total number of patients at risk for each sorted time point

```r
Y. <- apply(Y,2,sum)
```

### Nj. is matrix of number of failures up to sorted times with rows corresponding for each
  risk type and columns for sorted time points

```r
Nj. <- apply(N, c(1,3), sum)
```

### N.k is matrix of number of failures up to sorted times with rows corresponding for each
  group and columns for sorted time points
N.k <- apply(N, c(2,3), sum)

## N.N is matrix with number of failures up to each sorted time point
N.. <- apply(N, 3, sum)

## create array with cumulative CSH's, hazards, survival, and CIF's

times = s
ccsh <- array( dim=c(Is,(3*m) + 4,K) )
dimnames(ccsh) = list(NULL, c(rep(NA, (3*m)+4)), NULL)
dimnames(ccsh)[[2]][1] = "time"
dimnames(ccsh)[[2]][2:(m+1)] = paste("Haz",as.character(1:m),sep="")
dimnames(ccsh)[[2]][(m+2):(2*m+1)] = paste("Haz",as.character(1:m),sep="")
dimnames(ccsh)[[2]][(2*m+2):(3*m+1)] = paste("CI",as.character(1:m),sep="")
dimnames(ccsh)[[2]][(3*m+2)] = "hazsum"
dimnames(ccsh)[[2]][(3*m+3)] = "Hazsum"
dimnames(ccsh)[[2]][(3*m+4)] = "surv"

## Create m*K arrays for each of Z1, Z2, Z3
Z1 <- array( dim=c(m,K) )
Z2 <- array( dim=c(m,K) )
Z3 <- array( dim=c(m,K) )

## Compute cumulative CSH's, hazards, survival, and CIF's and put into array
## Fill in entries of the Z matrices that correspond to the Z-jk test processes estimates
## diff(N[j,k]) forms a vector with all of the entries of dN-jk(s) in formulas (great way to
compute x[(1+lag):n] = x[1:(n-lag)] where lag=1)
for ( j in 1:m ) {
    for ( k in 1:K ) {  
        Yk = Y[k,]
        Yk[Yk==Oj = 1
        #ccsh[,1,k] = times
        ccsh[,"time",k] = times
        ccsh[1,j+1,k] = 0
        ccsh[c(2:ls),j+1,k] = cumsum(diff(N[j,k])/Yk) # Cumulative CSH for risk j
        ccsh[1,m+1+j,k] = 0
        ccsh[c(2:ls),m+1+j,k] = diff(ccsh[,j+1,k]) # CSH for risk j
    }
}
for ( k in 1:K ) {  
    #ccsh[,2*m+2,k] = apply(ccsh[,((m+2):(2*m+1)),k],1,sum) # sum up all hazards
```r
#ccsh[,,(3*m+2),k] <- summary(survfit(Surv(ccsh[, "time", k],
ifelse((m==1), ccsh[,,(3*m+2),k] <- ccsh[,3,k], ccsh[,,(3*m+2),k] <- apply(ccsh[,((m+2):(2*m+1)),k],1,sum)) # sum up all hazards
#ccsh[,2*m+3,k] <- cumprod(1-ccsh[,2*m+3,k]) # survival fns for all risks pooled together
#ccsh[,"surv",k] <- cumprod(1-ccsh[,2*m+3,k]) # survival fns for all risks pooled together
ccsh[,,(3*m+4),k] <- cumprod(1-ccsh[,,(3*m+2),k]) # survival fns for all risks pooled together
#ccsh[,2*m+4,k] <- -log(ccsh[,2*m+3,k]) # this calculates sum of all cumulative hazards
#ccsh[,"Hazsum",k] <- -log(ccsh[,2*m+3,k]) # this calculates sum of all cumulative hazards
ccsh[,,(3*m+3),k] <- -log(ccsh[,,(3*m+4),k]) # this calculates sum of all cumulative hazards
#ccsh[,,(3*m+3),k] <- apply(ccsh[,,(2:(m+1)),k],1,sum) # this calculates sum of all cumulative hazards
}

for (j in 1:m) {
  for (k in 1:K) {
    ccsh[,2*m+j+1,k] = cumsum(c(1,ccsh[-1s,(3*m+4),k]) * ccsh[,,(m+1+j),k]) # CIF for risk j
    #ccsh[,3*m+j+1,k] = cumsum(c(1,ccsh[-1s,"Hazsum",k])*ccsh[,,(m+1+j),k]) # CIF for risk j = Hazsum * csh for risk j ???
  }
}

for (j in 1:m) {
  for (k in 1:K) {
    if (weight == 1) {
      Wjk = Y[k,j] * (Y./n)^rho
      ## Wjk = Y[k,j]
    } else if (weight == 2) {
      Yk = Y[k,k]
      Yk[Yk==0] = 1
      Sk = exp(-cumsum(diff(N,k[k],)/Y(k))
      #Sk = ccsh[,"surv",k]
      Wjk = Y[k,k] * (Sk)^r * (1-Sk)^r
    } else if (weight == 3) {
      Fj = exp(-cumsum(diff(N,k)/Y())) * cumsum(diff(N[,j])/Y.)
    }
  }
}
```
Wjk = Y * (1-Fj)·rho

Yk = Y[k,]; Yk[Yk==0] = 1; \# so we can divide by Yk in next step
Z1[j,k] <- sum( Wjk*(diff(N[j,k,])/Yk-diff(N..)/(m*Y)) )
Z2[j,k] <- sum( Wjk*(diff(N[j,k,])/Yk-diff(N[j,..])/Y) )
Z3[j,k] <- sum( Wjk*(diff(N[j,k,])/Yk-diff(N.k[k,])/m*Yk) )

# Find SI
mK <- m * K
SI <- array( dim=c(mK,mK) )
for ( jk in 1:mK ) {
  j <- (jk+K-1)%%K    \# x %% y integer division: 5 %% 2 is 1
  k <- (jk+K-1)%%K + 1 \# x %% y is modulus (x mod y): 5 %% 2 is 1
  for ( jlkl in 1:mK ) {
    jl <- (jlkl+K-1)%%K
    kl <- ((jlkl+K-1)%%K) + 1
    if (weight == 1) {
      Wjk = Y[k,] * (Y./n)'rho
      Wjlkl = Y[kl,] * (Y./n)'rho
      Wjk = Y[k,] \# Wjlkl = Y[kl,]
    } else if (weight == 2) {
      Yk = Y[k,]
      Yk[Yk==0] = 1
      Ykl = Y[kl,]
      Ykl[Ykl==0] = 1
      Sk = exp(-cumsum(diff(N.k[k,])/Yk))
      Skl = exp(-cumsum(diff(N.k[kl,])/Ykl))
      Wjk = Y[k,] * (Sk)'rho * (1-Sk)'r
      Wjlkl = Y[kl,] * (Skl)'rho * (1-Skl)'r
    } else if (weight == 3) {
      Fj = exp(-cumsum(diff(N..)/Y)) * cumsum(diff(N[j,..])/Y..)
      Fjl = exp(-cumsum(diff(N..)/Y)) * cumsum(diff(N[j,..])/Y..)
      Wjk = Y * (1-Fj)'rho
      Wjlkl = Y * (1-Fj)'rho
    }
  }
  Yk = Y[k,]; Yk[Yk==0] = 1;
\[ tmp = W_{jk} \cdot W_{jkl} \cdot \left( \frac{(j=j_1) \cdot (k=k_1)}{Y_{k-l}/Y_{.}} \right) \cdot \text{diff}(N_{.})/m \cdot Y_{.} \]

\[ S1[jk,jkl] = \text{sum}(tmp) \]

---

\# Find S2

\[ S2 \leftarrow \text{array}(\text{dim}=c(m,K,K)) \]

for (j in 1:m) {
  for (k in 1:K) {
    for (kl in 1:K) {
      if (weight = 1) {
        W_{jk} = Y_{[k,] \cdot (Y_{./n})^\rho}
        W_{jkl} = Y_{[kl,] \cdot (Y_{./n})^\rho}
        \# W_{jk} = Y_{[k,]}
        \# W_{jkl} = Y_{[kl,]}
      } else if (weight = 2) {
        Y_{k} = Y_{[k,]}'
        Y_{k}[Y_{k}==0] = 1
        Y_{kl} = Y_{[kl,]}
        Y_{kl}[Y_{kl}==0] = 1
        S_{k} = \exp(-\text{cumsum}(\text{diff}(N_{k[j,]})/Y_k))
        S_{kl} = \exp(-\text{cumsum}(\text{diff}(N_{k[kl,]})/Y_{kl}))
        \# S_{k} = \text{cosh}[\cdot \text{surv},k]
        \# S_{kl} = \text{cosh}[\cdot \text{surv},kl]
        W_{jk} = Y_{[k,] \cdot (S_{k})^\rho \cdot (1-S_{k})^r}
        W_{jkl} = Y_{[kl,] \cdot (S_{kl})^\rho \cdot (1-S_{kl})^r}
      } else if (weight = 3) {
        F_{j} = \exp(-\text{cumsum}(\text{diff}(N_{[j,]})/Y_{.})) \cdot \text{cumsum}(\text{diff}(N_{j[j,]})/Y_{.})
        W_{jk} = Y_{.} \cdot (1-F_{j})^\rho
        W_{jkl} = Y_{.} \cdot (1-F_{j})^\rho
      } \]

    \} \]

  } \]

} \]

---

\# Find S3

\[ S3 \leftarrow \text{array}(\text{dim}=c(m,m,K)) \]

for (j in 1:m) {
  for (j1 in 1:m) {
    \} \]

} \]
for (k in 1:K) {
  if (weight == 1) {
    Wjk = Y[k,] * (Y./n)^rho
    Wjlk = Y[k,] * (Y./n)^rho
    # Wjk = Y[k,]
    # Wjlk = Y[k,]
  } else if (weight == 2) {
    Yk = Y[k,]
    Yk[Yk==0] = 1
    Sk = exp(-cumsum(diff(N.k[k,])/Yk))
    Wjk = Y[k,] * (Sk)^rho * (1-Sk)^r
    Wjlk = Y[k,] * (Sk)^rho * (1-Sk)^r
  } else if (weight == 3) {
    Fj = exp(-cumsum(diff(N.j[j,])/Y.)) * cumsum(diff(N.j[j,])/Y.)
    Fj1 = exp(-cumsum(diff(N.j[j,])/Y.)) * cumsum(diff(N.j[j,])/Y.)
    Wjk = Y. * (1-Fj)^rho
    Wjlk = Y. * (1-Fj)^rho
  }
  Yk = Y[k,]; Yk[Yk==0] = 1;
  tmp = (((j==j1)-1/m) * Wjk*Wjlk /Yk * diff(N.k[k,])/m/Yk
  S3[j,j1,k] = sum(tmp)
}

Find T1
Z <- as.vector(t(Z1))[-mK]  # using index of [-mK] excludes the mKth entry from the vector (this is the last entry in the vector)
S <- S1[-mK,-mK]  # using index of [-mK,-mK] excludes the mKth row and mKth column from the matrix
T1 <- as.numeric( Z %*% ginv(S) %*% Z )  # T1 <- sum(Z*solve(S,Z))

Find T2
if(K > 1) {
  T2stats <- T2pvalues <- rep(NA, m)
  T2 <- 0;
  for (j in 1:m) {
    Z <- Z2[j,-K]
S <- S2[j, ,]
S <- S[-K,-K]

## T2 <- T2 + sum(Z*solve(S,Z))
T2 <- T2 + as.numeric(Z %*% ginv(S) %*% Z)
T2stats[j] <- as.numeric(Z %*% ginv(S) %*% Z)

# T2pvalues[j] = pchisq(T2stats[j], (K-1), lower.tail=F)
T2pvalues[j] = ifelse(pchisq(T2stats[j], (K-1), lower.tail=F) < 0.0001, "<0.0001",
"%6.4f", 1-pchisq(T2stats[j], (K-1), lower.tail=T))

}

if (m > 1) {
    # Find T3
    T3stats <- T3pvalues <- rep(NA.K)
    T3 <- 0;
    for (k in 1:K) {
        Z <- Z3[-m,k]
        S <- S3[,,k]
        S <- S[-m,-m]
        ## T3 <- T3 + sum(Z*solve(S,Z))
        T3 <- T3 + as.numeric(Z %*% ginv(S) %*% Z)
        T3stats[k] <- as.numeric(Z %*% ginv(S) %*% Z)
        # T3pvalues[k] = pchisq(T3stats[k], (m-1), lower.tail=F)
        T3pvalues[k] = ifelse(pchisq(T3stats[k], (m-1), lower.tail=F) < 0.0001, "<0.0001",
"%6.4f", 1-pchisq(T3stats[k], (m-1), lower.tail=T))
    }
}

if (K > 1) {
    # pvalueT1 = pchisq(T1, m*K-1, lower.tail=F)
    pvalueT1 = ifelse(pchisq(T1, m*K-1, lower.tail=F) < 0.0001, "<0.0001",
"%6.4f", pchisq(T1, m*K-1, lower.tail=F))
    # pvalueT2 = pchisq(T2, m*(K-1), lower.tail=F)
    pvalueT2 = ifelse(pchisq(T2, m*(K-1), lower.tail=F) < 0.0001, "<0.0001",
"%6.4f", pchisq(T2, m*(K-1), lower.tail=F))
    if (m > 1) {
        # pvalueT3 = pchisq(T3, (m-1)*K, lower.tail=F)
        pvalueT3 = ifelse(pchisq(T3, (m-1)*K, lower.tail=F) < 0.0001, "<0.0001",
"%6.4f", pchisq(T3, (m-1)*K, lower.tail=F))
    }
    # else if (m == 1) {
    # list(ccsh=ccsh, T2=T2, df=m*(K-1), pvalues=pvalueT2)
    #}
else if (K == 1) {
    # pvalueT3 = pchisq(T3, (m-1)*K, lower.tail=F)
    pvalueT3 = ifelse(pchisq(T3, (m-1)*K, lower.tail=F) < 0.0001, "<0.0001", sprintf("%6.4f", pchisq(T3, (m-1)*K, lower.tail=F)))
}

if (plot="T") {
    ## Construct plot of either cuminc curves or ccsf curves
    ## est_array is a multidimensional array composed of m (number of risks) matrices with
dimensions rows=timepoints and columns=number of groups

times = ccsh[,"time",1]
max_time = max(ccsh[,"time",1])
est_array = array(dim=c(length(ccsh[,"time",1]),K,m))
legendtext <- matrix(NA, nrow=K, ncol=m)
bank = c(l,4,6,3,5,2,5,1,3)
bank = bank[1:m*K]
linetypes <- matrix(bank, nrow=K, ncol=K)
for (j in 1:m) {
    for (k in 1:K) {
        legendtext[k,j] = paste("DSCAT",j,"",covar_level="",k)
    }
}
if (type == "ccsh") {
    # for (k in 1:K) {
    #    est-ae[,k] = ccsh,"Hazl",k]
    #    est-loe[,k] = ccsh,"Haz2",k]
    #    est-oth[,k] = ccsh,"Haz3",k]
    # }
    for (j in 1:m) {
        for (k in 1:K) {
            est_array[,k,j] = ccsh[,j+1,k]
        }
    }
    ylab = "Cumulative_Hazard_Rate"
} else if (type="cuminc") {
    # for (k in 1:K) {
    #    est-ae[,k] = ccsh,"CI1",k]
    #    est-loe[,k] = ccsh,"CI2",k]
    #    est-oth[,k] = ccsh,"CI3",k]
}
for (j in 1:m) {
  for (k in 1:K) {
    est.array[,k,j] = ccsh[,2*m+j+1,k]
  }
}

ylab = "Cumulative_Incidence_Probability"

max.plot = 0.05 + round(max(est.array),2)
est.array.firstplot = est.array[,1,1]
linetypesnew = linetypes
linetypefirst = linetypes[1,1]
#lrdim = nrow(linetypesnew)*ncol(linetypes)
#linetypesvector = rep(NA, length(as.vector(linetypesnew)))
linetypesvector = as.vector(linetypesnew)

if (dscatplot > 0) {
est.array.firstplot = est.array[,1,dscatplot]
linetypefirst = linetypes[1,dscatplot]
est.array = est.array[,dscatplot]
max.plot = 0.05 + round(max(est.array),2)
legendtext = legendtext[,dscatplot]
linetypesnew = linetypes[,dscatplot]
#linetypesvector = rep(NA, length(as.vector(linetypesnew)))
linetypesvector = as.vector(linetypesnew)
}

if (groupplot > 0) {
est.array.firstplot = est.array[,groupplot,1]
linetypefirst = linetypes[groupplot,1]
est.array = est.array[,groupplot,]
max.plot = 0.05 + round(max(est.array),2)
legendtext = legendtext[groupplot,]
linetypesnew = linetypes[groupplot,]
#linetypesvector = rep(NA, length(as.vector(linetypesnew)))
linetypesvector = as.vector(linetypesnew)
}

plot (ccsh[,"time",1],
est.array.firstplot,
type="s",
ylim=c(0,max.plot),
lt=linetypefirst,
xlab="Duration_(Weeks)",
ylab=ylab,
main=datatext, sub=vartext)
if (dscatplot > 0) {
  for (j in dscatplot) {
    for (k in 1:K) {
      lines (ccsh[, "time", 1], est_array[, k], type="s", lty=linetypes[k, j])
    }
  }
}

if (groupplot > 0) {
  for (j in 1:m) {
    for (k in groupplot) {
      lines (ccsh[, "time", 1], est_array[, j], type="s", lty=linetypes[k, j])
    }
  }
}

if (dscatplot == 0 & groupplot == 0) {
  for (j in 1:m) {
    for (k in 1:K) {
      lines (ccsh[, "time", 1], est_array[, k, j], type="s", lty=linetypes[k, j])
    }
  }
}

if (legend="T") {
  legend ("topleft", inset = 0.05, legend=text)
  #lty=c((1:(m*K)))
  #lty=as.vector(linetypesnew)[1]:as.vector(linetypesnew)[ldim])
  lty=as.vector(linetypesnew))
}

if (K == 1 & m > 1) {
  text (x=max.time,
y = max_plot - 0.02, # MDD, DPNP, FM,
# labels = paste("TKC test 1: p-value ", pvalueT1, ",TKC test 2: p-value ",
# pvalueT2, ",TKC test 3: p-value ", pvalueT3),
labels = paste("TKC test 1: p-value ", pvalueT1, ",TKC test 2: p-value ",
# pvalueT2, ",TKC test 3: p-value ", pvalueT3),
cex = 1.1, # MDD (cuminc, ccsh), DPNP (cuminc, ccsh), FM (cuminc, ccsh),
pos = 2
}
if (K > 1 & m == 1) {
  text (x = max_time, 
  y = max_plot - 0.02, # MDD, DPNP, FM, 
  labels = paste("TKC test 1: p-value ", pvalueT1, ",TKC test 2: p-value ",
# pvalueT2, ",TKC test 3: p-value ", pvalueT3),
cex = 1.1, # MDD (cuminc, ccsh), DPNP (cuminc, ccsh), FM (cuminc, ccsh),
pos = 2
}
else if (K > 1 & m > 1) {
  text (x = max_time, 
  y = max_plot - 0.02, # MDD, DPNP, FM, 
  labels = paste("TKC test 1: p-value ", pvalueT1, ",TKC test 2: p-value ",
# pvalueT2, ",TKC test 3: p-value ", pvalueT3),
cex = 1.1, # MDD (cuminc, ccsh), DPNP (cuminc, ccsh), FM (cuminc, ccsh),
pos = 2
}
else if (dscatplot > 0 & groupplot > 0) {
  text (x = max_time, 
  y = max_plot - 0.02, # MDD, DPNP, FM, 
  labels = paste("TKC test 1: p-value ", pvalueT1, ",TKC test 2: p-value ",
# pvalueT2, ",TKC test 3: p-value ", pvalueT3),
  # cex = 1.1, # MDD (cuminc, ccsh), DPNP (cuminc, ccsh), FM (cuminc, ccsh),
  # pos = 2
  # }
  # }
if (K > 1) {
  if (m > 1) {
    list (ccsh = ccsh, teststats = c(T1, T2, T3), df = c(m * K - 1, m * (K - 1), (m - 1) * K), pvalues =
      (pvalueT1, pvalueT2, pvalueT3), T2stats = T2stats, T2values = T2values, 
      T3stats = T3stats, T3values = T3values, times = times, est_array = est_array, 
      legendtext = legendtext, linetypes = linetypes)
  } else if (m == 1) {
A.6: GAD Treatment Assignment Plot Generator

# 1. Create plot of all risks and all levels of treatment (use same colors for risk, just different line types for TRT and PLA
# 2. Look more closely at the significance of T2 and T3:
# A. Create separate plots of individual DC risk with each group level (Looking at T2)
# B. Create separate plots of individual Group levels with each DC risk (Looking at T3)
## Graph notes
## If doing colored lines, uncomment the col parameters in each plot call and have each pair of curves have lty=1 and lty=2
## If doing black-white plots, comment out the col parameters and use lty = 1,2,3,4,5,6 for each curve respectively

```r
list( ccsh=ccsh, teststats=c(T2), df=c(m*(K-1)), pvalues=pvalueT2, times=times,
est_array=est_array, legendtext=legendtext, linetypes=linetypes)
}
}
else if (K==1) {
list( ccsh=ccsh, teststats=c(T3), df=c((m-1)*K), pvalues=pvalueT3, times=times,
est_array=est_array, legendtext=legendtext, linetypes=linetypes)
}
}
else if (plot=="F") {
  if (K > 1) {
    if (m > 1) {
      list( ccsh=ccsh, teststats=c(T1,T2,T3), df=c(m*K-1, m*(K-1), (m-1)*K), pvalues=c(pvalueT1, pvalueT2, pvalueT3), T2stats=T2stats, T2pvalues=T2pvalues,
        T3stats=T3stats, T3pvalues=T3pvalues)
    }
    else if (m==1) {
      list( ccsh=ccsh, teststats=c(T2), df=c(m*(K-1)), pvalues=pvalueT2)
    }
  }
  else if (K==1) {
    list( ccsh=ccsh, teststats=c(T3), df=c((m-1)*K), pvalues=pvalueT3)
  }
}
```
source("R_funcs/FNweight.weights.R")

X = list(gad_dis$visit_week[gad_dis$strt_ind=="PLA"], gad_dis$visit_week[gad_dis$strt_ind=="TRT" ])

d = list(gad_dis$dscat[gad_dis$strt_ind=="PLA"], gad_dis$dscat[gad_dis$strt_ind=="TRT" ])

gad_tkc$trt_ind = weight.test(X, d, rho=0, r=0, weight=1, t=Inf, type="ccsh", plot="T", legend="T")

path="/home/eric/Dropbox/diss_latex/ch4/figures/"

#pdf(file=" /home/eric/Dropbox/diss_latex/interview-presentation/figures/GAD-CCSH-TRT-ind.pdf")

pdf(file=paste(path,"GAD_CCSH_TRT_ind.pdf",sep=''))

par(mar=c(5,4,2,2)+0.1)

plot(gad_tkc$trt_ind$times,

gad_tkc$trt_ind$est[1,1],
type="s",
las=1,
cex.lab=1.2,
cex.axis=1.1,
ylim=0.05 + round(max(gad_tkc$trt_ind$est.array),2)),
#lty=gad_tkc$nogroups$linetypes[1],
#col="blue",
lwd=2,
lty=1,
xlab="Duration... (Weeks)",
ylab="Cumulative.CSH"
)

lines(gad_tkc$trt_ind$times,

gad_tkc$trt_ind$est.array[,2,1],
type="s",
#lty=gad_tkc$trt_ind$linetypes[2],
#col="blue",
lwd=2,
lty=2
)

lines(gad_tkc$trt_ind$times,

gad_tkc$trt_ind$est.array[,1,2],
type="s",
```r
# graphic:

# line types
lty=gad_tkc_trt_ind$linetypes[2],
# color = "red",
lwd=2,
lty=3

lines (gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,2,2],
type="s",
# lty=gad_tkc_trt_ind$linetypes[2],
# col = "red",
lwd=2,
lty=4)

lines (gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,1,3],
type="s",
# lty=gad_tkc_trt_ind$linetypes[2],
# col = "green",
lwd=2,
lty=5)

lines (gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,2,3],
type="s",
# lty=gad_tkc_trt_ind$linetypes[2],
# col = "green",
lwd=2,
lty=6)

legend ( "topleft",
inset = 0.01,
legend=c("DC due to AE: PLA", "DC due to AE: TRT", "DC due to LOE: PLA", "DC due to LOE: TRT", "DC due to OTH: PLA", "DC due to OTH: TRT"),
lty=c(1,2,3,4,5,6),
lwd=c(2,2,2,2,2,2),
# cex=c(1.2,1.2,1.2),
# col=c("blue", "blue", "red", "red", "green", "green")
# lty=c(as.vector(linetypesnew)[1]:as.vector(linetypesnew)[ltdim])
)

text(x=max(gad_tkc_trt_ind$times)-5,
y=max(gad_tkc_trt_ind$est_array)+0.045,
adj=0,
cex=1.2,
as.expression(substitute(T[l]= unknown, list(unknown=gad_tkc_trt_ind$teststats[l])))
```
dev.off()

### Plot of DSCAT=1 and both treatment levels

```r
pdf(file=paste(path,"GAD_CCSH_TRT_ind-DSCAT1.pdf",sep=' '))
#pdf(file="/home/eric/Dropbox/diss_latex/interview_presentation/figures/GAD_CCSH_TRT_ind_DSCAT1.pdf")
par(mar=c(5,4,2,2)+0.1)
plot(gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,1,1],
type="s",
```
las=1,
cex.lab=1.2,
cex.axis=1.1,
ylim=c(0,0.05 + round(max(gad_tkc_trt_ind$est.array),2)),
#lty=gad_tkc_trt$linetypes[1],
#col="blue",
lwd=2,
lty=1,
xlab="Duration_(Weeks)",
ylab="Cumulative_CSH"
)
lines(gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est.array[,2,1],
type="s",
#lty=gad_tkc_trt$linetypes[2],
#col="blue",
lwd=2,
lty=2
)
legend("topleft",
 inset=0.01,
legend=c("DC_due_to_AE:_PLA", "DC_due_to_AE:_TRT"),
lty=c(1,2),
lwd=c(2,2),
#cex=c(1.2,1.2,1.2),
#col=c("blue", "blue")
#lty=c(as.vector(linetypesnew)[1]:as.vector(linetypesnew)[ltdim]))
)
text(x=max(gad_tkc_trt_ind$times)-2.5,
y=max(gad_tkc_trt_ind$est.array)+0.045,
adj=0,
cex=1.2,
as.expression(substitute(T[2]==unknown, list(unknown=gad_tkc_trt_ind$T2stats[1])))
)
text(x=max(gad_tkc_trt_ind$times)-2.5,
y=max(gad_tkc_trt_ind$est.array)+0.025,
adj=0,
cex=1.2,
as.expression(substitute(pvalue:unknown, list(unknown=gad_tkc_trt_ind$T2pvalues[1])))
)
dev.off()

## Plot of DSCAT=2 and both treatment levels
\begin{verbatim}
par(mar=c(5,4,2,2)+0.1)
plot (gad.tkc_trt_ind$times, gad_tkc_trt_ind$est.array[,1,2], type="s", las=1, cex.lab=1.2, cex.axis=1.1, ylim=c(0,0.05 + round(max(gad_tkc_trt_ind$est.array),2)), #lty=gad_tkc_nogroups$linetypes[1], #col="red", lwd=2, lty=3, xlab="Duration (Weeks)", ylab="Cumulative CSH"
)
lines (gad.tkc_trt_ind$times, gad_tkc_trt_ind$est.array[,2,2], type="s", #lty=gad_tkc_trt_ind$linetypes[2], #col="red", lwd=2, lty=4
)
legend ( "topleft", inset = 0.01, legend=c("DC_due_to_LOE:_PLA", "DC_due_to_LOE:_TRT"), lty=c(3,4), lwd=c(2,2), #cex=c(1.2,1.8,1.2), #col=c("red", "red") #lty=c(as.vector(linetypesnew)[1]:as.vector(linetypesnew)[ldim])
)
text(x=max(gad.tkc_trt_ind$times)-2.5, y=max(gad.tkc_trt_ind$est.array)+0.045, adj=0, cex=1.2, as.expression(substitute(T[2]==unknown, list(unknown=gad.tkc_trt_ind$T2stats[2])))
)
text(x=max(gad.tkc_trt_ind$times)-2.5,
\end{verbatim}
```r
y = max(gad_tkc_trt_ind$est_array) + 0.025,
adj = 0,
cex = 1.2,

as.expression(substitute(pvalue = unknown, list(unknown = gad_tkc_trt_ind$T2pvalues[2])))

dev.off()

########################################################
## Plot of DSCAT=3 and both treatment levels
########################################################

df <- paste(path, "GAD_CCSH_TRT_ind_DSCAT3.pdf", sep = "")

par(mar = c(5, 4, 2, 2) + 0.1)

plot(gad_tkc_trt_ind$times, gad_tkc_trt_ind$est_array[, 1, 3],
type = "s",
las = 1,
cex.lab = 1.2,
cex.axis = 1.1,
ylim = c(0, 0.05 + round(max(gad_tkc_trt_ind$est_array), 2)),
#lty = gad-tkc-nogroups$linetypes[1],
#col = "green",
lwd = 2,
lty = 5,

xlab = "Duration (Weeks)",
ylab = "Cumulative CSH"
)

lines(gad_tkc_trt_ind$times, gad_tkc_trt_ind$est_array[, 2, 3],
type = "s",
#lty = gad_tkc_trt_ind$est_array[, 2],
#col = "green",
lwd = 2,
lty = 6
)

legend("topleft",
inset = 0.01,
legend = c("DC_due_to_OTH: PLA", "DC_due_to_OTH: TRT"),
lty = c(5, 6),
lwd = c(2, 2),
#cex = c(1.2, 1.2, 1.2),
```

## Plot of trt_ind="PLA" and all three risks

```r
pdf(file=paste(path,"GAD_CCSH_TRT_ind-PLA.pdf",sep=''))
par(mar=c(5,4,2,2)+0.1)
plot(gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,1,1],
type="s",
las=1,
cex.lab=1.2,
cex.axis=1.1,
ylim=c(0,0.05 + round(max(gad_tkc_trt_ind$est_array),2)),
#lty=gad_tkc_nogroups$linetypes[1],
#col="red",
lwd=2,
lty=1,
xlab="Duration (Weeks)",
ylab="Cumulative CSH"
)
lines(gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,1,2],
type="s",
#lty=gad_tkc_trt_ind$linetypes[2],
```
```R
#col="blue",
lwd=2,
lty=3
)
lines (gad_tkc_trt_ind$times,
gad_tkc_trt_ind$est_array[,1,3],
type="s",
#lty=gad_tkc_trt_ind$linetypes[2],
#col="green",
lwd=2,
lty=5
)
legend ( "topleft",
inset = 0.01,
legend=c("DC_due_to_AE: PLA", "DC_due_to_LOE: PLA", "DC_due_to_OTH: PLA"),
lty=c(1,3,5),
lwd=c(2,2,2),
#cex=c(1.2,1.2,1.2),
#col=c("red", "blue", "green")
#lty=c(as.vector(linetypesnew)[1]:as.vector(linetypesnew)[ltdim]))
)
text(x=max(gad_tkc_trt_ind$times)-2.5,
y=max(gad_tkc_trt_ind$est_array)+0.045,
adj=0,
cex=1.2,
expression(substitute(T[3]==unknown, list(unknown=gad_tkc_trt_ind$T3stats[1])))
)
text(x=max(gad_tkc_trt_ind$times)-2.5,
y=max(gad_tkc_trt_ind$est_array)+0.025,
adj=0,
cex=1.2,
expression(substitute(pvalue:unknown, list(unknown=gad_tkc_trt_ind$T3pvalues[1])))
)
dev.off()
```

### Plot of trt.ind="TRT" and all three risks

```
pdf(file=paste(path,"GAD_CCSH_TRT_ind-TRT.pdf",sep=""))
```

```
pdf(file="/home/eric/Dropbox/diss_latex/interview_presentation/figures/GAD_CCSH_TRT_ind-TRT.pdf")
```

```
par(mar=c(5,4,2,2)+0.1)
plot (gad_tkc_trt_ind$times,
```
```r
gad_tkc_trt_ind$est_array[,2,1],
  type="s",
  las=1,
  cex.lab=1.2,
  cex.axis=1.1,
ylim=c(0,0.05 + round(max(gad_tkc_trt_ind$est_array),2)),
  #lty=gad_tkc_mgroups$linetypes[1],
  #col="red",
  lwd=2,
  lty=1,
  xlab="Duration (Weeks)",
ylab="Cumulative_CSH"
)
lines(gad_tkc_trt_ind$times,
  gad_tkc_trt_ind$est_array[,2,2],
  type="s",
  #lty=gad_tkc_trt_ind$linetypes[2],
  #col="blue",
  lwd=2,
  lty=4
)
lines(gad_tkc_trt_ind$times,
  gad_tkc_trt_ind$est_array[,2,3],
  type="s",
  #lty=gad_tkc_trt_ind$linetypes[2],
  #col="green",
  lwd=2,
  lty=6
)
legend("topleft",
  inset = 0.01,
  legend=c("DC_due_to_AE:TRT", "DC_due_to_LOE:TRT", "DC_due_to_OTH:TRT"),
  #lty=c(2,4,6),
  lty=c(1,3,5),
  lwd=c(2,2,2),
  cex=c(1.2,1.2,1.2),
  #col=c("red", "blue", "green")
  #lty=c(as.vector(linetypesnew)[1]:as.vector(linetypesnew)[ldim]))
)
text(x=max(gad_tkc_trt_ind$times) - 2.5,
  y=max(gad_tkc_trt_ind$est_array) + 0.045,
```
adj=0,
cex=1.2,
as.expression(substitute(T[3]==unknown, list(unknown=gad_tkc_trt_ind$T3stats[2])))
)
text(x=max(gad_tkc_trt_ind$times)-2.5,
y=max(gad_tkc_trt_ind$est_array)+0.025,
adj=0,
cex=1.2,
as.expression(substitute(pvalue:unknown, list(unknown=gad_tkc_trt_ind$T3pvalues[2])))
)
dev.off()

A.7: Fine and Gray Regression Model Selection Program

crr.cont_model_selection <- function(dataset, datatext, risk, var) {

## Step 1: Compare null model with univariate models to get significant predictors in
## univariate

dev <- pval <- varname <- index <- loglik.null <- loglik <- shr <- c(rep(NA, length(var)))
risk = c(rep(risk, length(var)))
 df = c(rep(1, length(var)))
disease = c(rep(datatext, length(var)))

for(j in 1:length(var)) {
  varname[j] = var[j]
  index[j] = j
  #fit = crr(dataset$visit_week, dataset$dscat_kly, dataset[var[j]], failcode=risk)
  fit = crr(dataset$visit_week, dataset$dscat, dataset[var[j]], failcode=risk)
  #print(summary(fit))
  shr[j] = round(summary.crr(fit, digits=4)$coef[, "exp(coef)"] ,4)
  loglik.null[j] = fit$loglik.null
  #print(loglik.null[j])
  loglik[j] = fit$loglik
  #print(loglik[j])
  dev[j] = round(2*(loglik[j] - loglik.null[j]),4)
  pval[j] = round(pchisq(dev[j], df=1, lower.tail=F),4)
}
results.step1 = data.frame(disease, risk, varname, index, shr, loglik.null, loglik, dev, df,
pval)
varsig.step1 = as.vector(results.step1$varname[results.step1$pval<0.1])
varsig.step1.indexes = results.step1$index[results.step1$pval<0.1]
cat("Step.1-Analysis\n","Significant\nvariables_from_univariate-analysis\n")
print(varsig.step1)
## Step 2: Fit full model with all sig vars from step 1 and try leaving out one variable at a time to see which vars are sig still

```r
disease_step2 = rep(datatext, times=length(varsig_step1_indexes))

### risk is way too long of a vector so leave it out of data frame formulations
risk_step2 = rep(risk, times=length(varsig_step1_indexes))

index_step2 = seq(along=varsig_step1_indexes)
```

```r
fit_full_step1 = crr(dataset$visit-week, dataset$dscat, dataset[varsig[c(varsig_step1_indexes)]], failcode = risk)

loglik_step1 = fit_full_step1$loglik

loglik_step1_vec = rep(round(fit_full_step1$loglik, 4), times=length(varsig_step1_indexes))
```

```r
# i = 1
for(j in 1:length(varsig_step1_indexes)) {
  varname[j] = varsig_step1[j]
  varpos = which(var==varsig_step1[j])
  varsig_new_indexes = varsig_step1_indexes[-j]
  index[j] = j
  fit = crr(dataset$visit-week, dataset$dscat, dataset[varsig[c(varsig_new_indexes)]], failcode = risk)
  shr[j] = round(summary.crr(fit, digits=4)$coef[, "exp(coef)"]/4)
  loglik[j] = fit$loglik
  dev[j] = round(2*(loglik_step1 - loglik[j]), 4)
  pval[j] = pchisq(dev[j], df=1, lower.tail=F)
```
```r
#cat("pval for loop index ",j," is ",pval[j],"\n")

results_step2 = data.frame(varname, varsig_step1_indexes, loglik_step1_vec, loglik, dev, pval)
varsig_step2 = as.vector(results_step2$varname[results_step2$pval<0.1])
varsig_step2_indexes = results_step2$varsig_step1_indexes[results_step2$pval<0.1]
cat("Significant variables from step 2\n")
print(varsig_step2)
cat("\n")
#print(varsig_step2_indexes)
#cat("\n\n")

## Step 3: Try variables that were discarded in step 1 into full model from step 2

vars_to_try = vars[-varsig_step2_indexes]
vars_to_try_indexes = rep(NA, length(vars_to_try))
#vars_to_try_indexes = rep(NA, length(var)-length(varsig_step2_indexes))
dev <- pval <- varname <- loglik <- shr <- c(rep(NA, length(vars_to_try)))

fit_full_step2 = crr(dataset$visit_week, dataset$dscat, dataset[var[c(varsig_step2_indexes)]], failcode=risk)
loglik_step2 = fit_full_step2$loglik
loglik_step2_vec = rep(round(fit_full_step2$loglik,4), times=length(vars_to_try))

for (j in 1:length(vars_to_try)) {
  #cat("loop index for step 3 = ",j,"\n")
  varname[j] = vars_to_try[j]
  vars_to_try_indexes[j] = which(var==vars_to_try[j])
  fit = crr(dataset$visit_week, dataset$dscat, dataset[var[c(vars_to_try_indexes[j]), varsig_step2_indexes]]), failcode=risk)
  #print(summary(fit))
  loglik[j] = fit$loglik
  #cat("loglik for loop index ",j," is ",loglik[j],"\n")
  #dev[j] = round(-2*(loglik_step1 - fit$loglik),4)
  dev[j] = round(2*(loglik[j] - loglik_step2),4)
  #cat("dev for loop index ",j," is ",dev[j],"\n")
  pval[j] = pchisq(dev[j], df=1, lower.tail=F)
  #cat("pval for loop index ",j," is ",pval[j],"\n")
}

results_step3 = data.frame(varname, vars_to_try_indexes, loglik_step2_vec, loglik, dev, pval)
varsig_step3 = as.vector(results_step3$varname[results_step3$pval<0.1])
```
varsig_step3.indexes = results_step3$vars_to_try_indexes[results_step3$pval < 0.1]
cat("Significant-variables-from-step-3\n")
print(varsig_step3)
cat("\n")
#print(varsig_step3.indexes)
#cat("\n\n")

## Step 4: Check to make sure no term in the model can be omitted without significantly increasing the value of $-2 \log L$

fit_full_step3 = crr(dataset$visit.week, dataset$dsat, dataset[, c(varsig_step2.indexes, varsig_step3.indexes)], failcode=risk)
loglik_step3 = fit_full_step3$loglik

loglik_step3.vec = rep(round(fit_full_step3$loglik, 4), times=length(c(varsig_step2.indexes, varsig_step3.indexes)))
dev <- pval <- varname <- index <- loglik <- shr <- c(rep(NA, length(c(varsig_step2.indexes, varsig_step3.indexes))))

varstochoose = c(varsig_step2, varsig_step3)

varstochoose_indexes = c(varsig_step2.indexes, varsig_step3.indexes)

for(j in 1:length(c(varsig_step2.indexes, varsig_step3.indexes))) {
  #cat("loop index for step 4 = ",j,"\n")
  varname[j] = varstochoose[j]
  #varpos = which(var==varsig_step1[j])
  varsig_new.indexes = varstochoose_indexes[-j]
  #index[j] = j
  fit = crr(dataset$visit.week, dataset$dsat, dataset[, c(varsig_new.indexes)], failcode = risk)
  #shr[j] = round(summary.crr(fit, digits=4)$coef,"exp(coef)",[4)
  loglik[j] = fit$loglik
  #cat("loglik for loop index ",j,"\n")
  #dev[j] = round(-2*(loglik_step3 - fit$loglik), 4)
  dev[j] = round(2*(loglik_step3 - loglik[j]), 4)
  #cat("dev for loop index ",j,"\n")
  pval[j] = pchisq(dev[j], df=1, lower.tail=F)
  #cat("pval for loop index ",j,"\n")
}

results_step4 = data.frame(varname, varstochoose_indexes, loglik_step3.vec, loglik, dev, pval)

varsig_step4 = as.vector(results_step4$varname[results_step4$pval < 0.1])

varsig_step4.indexes = results_step4$varstochoose_indexes[results_step4$pval < 0.1]
cat("Significant-variables-from-step-4\n")
print(varsig_step4)
## Step 5: Return data frame with parameters from final model

```r
fit_final = crr(dataset$visit_week, dataset$dscat, dataset[vars[c(varsig_step4_indexes)]], failcode=risk)
fit_final_vars = varsig_step4
fit_final_vars_indexes = varsig_step4_indexes
return(fit_final, fit_final_vars, fit_final_vars_indexes)
```

```r
# print(varsig_step2_indexes)
# cat("\n\n")
# fit_end = crr(dataset$visit_week, dataset$dscat, dataset[vars[c(varsig_step3_indexes, varsig_step2_indexes)]], failcode=risk)
# print(fit_end)
# return(data.frame(varname, loglik_step2, loglik, dev, pval))
# return (data.frame(disease, risk, varname, indexes, varsig, varsignew, shr, loglik.null, loglik, dev, df, pval))
# list(results = results)
```

### A.8: Fine and Gray Regression Model Selection Runner Program

```r
source("R_funcs/crr_cont-model_selection.R")
source("R_funcs/crr_results.R")
path="/home/eric/Dropbox/diss_latex/ch4/tables/"

#vars = c("ageyr", "pctchange_stand", "mxsevp_stand", "nsevp", "duration", "cau_ind", "country_cat_l", "sex_male", "trt_TRT")
vars = c("ageyr", "pctchange_stand", "mxsevp_stand", "duration", "cau_ind", "country_cat_l", "sex_male", "trt_TRT")
#vars = c("ageyr", "pctchange_stand", "mxsevp_stand", "duration", "country_cat_l", "sex_male", "trt_TRT")

## MDD
mdd_crr_1_results = crr_cont-model_selection(mdd.dis, "mdd", risk=1, vars)
mdd_crr_risk1_results = crr_results(mdd_crr_1_results$fit_final, "MDD", risk=1, mdd_crr_1_results$fit_final_vars)
mdd_crr_risk1_nrow = nrow(mdd_crr_risk1_results)

mdd_crr_2_results = crr_cont-model_selection(mdd.dis, "mdd", risk=2, vars)
mdd_crr_risk2_results = crr_results(mdd_crr_2_results$fit_final, "MDD", risk=2, mdd_crr_2_results$fit_final_vars)
```
mdd.err.risk2.nrow = nrow(mdd.err.risk2.results)

mdd.err.risk3.results = crr.cont.model_selection(mdd.dis, "mdd", risk=3, vars)
mdd.err.risk3.nrow = nrow(mdd.err.risk3.results)

## GAD

gad.err.risk1.results = crr.cont.model_selection(gad.dis, "gad", risk=1, vars)
gad.err.risk1.results = crr.results(gad.err.risk1.results$fit.final,"GAD",risk=1, gad.err.risk1.results$fit.final.vars)
gad.err.risk1.nrow = nrow(gad.err.risk1.results)

gad.err.risk2.results = crr.cont.model_selection(gad.dis, "gad", risk=2, vars)
gad.err.risk2.results = crr.results(gad.err.risk2.results$fit.final,"GAD",risk=2, gad.err.risk2.results$fit.final.vars)
gad.err.risk2.nrow = nrow(gad.err.risk2.results)

gad.err.risk3.results = crr.cont.model_selection(gad.dis, "gad", risk=3, vars)
gad.err.risk3.nrow = nrow(gad.err.risk3.results)

## DPNP

dpnp.err.risk1.results = crr.cont.model_selection(dpnp.dis, "dpnp", risk=1, vars)
dpnp.err.risk1.results = crr.results(dpnp.err.risk1.results$fit.final,"DPNP",risk=1, dpnp.err.risk1.results$fit.final.vars)
dpnp.err.risk1.nrow = nrow(dpnp.err.risk1.results)

dpnp.err.risk2.results = crr.cont.model_selection(dpnp.dis, "dpnp", risk=2, vars)
dpnp.err.risk2.results = crr.results(dpnp.err.risk2.results$fit.final,"DPNP",risk=2, dpnp.err.risk2.results$fit.final.vars)
dpnp.err.risk2.nrow = nrow(dpnp.err.risk2.results)

dpnp.err.risk3.results = crr.cont.model_selection(dpnp.dis, "dpnp", risk=3, vars)
dpnp.err.risk3.nrow = nrow(dpnp.err.risk3.results)

## FM

fibro.err.risk1.results = crr.cont.model_selection(fibro.dis, "fibro", risk=1, vars)
fibro.err.risk1.results = crr.results(fibro.err.risk1.results$fit.final,"FM",risk=1, fibro.err.risk1.results$fit.final.vars)
fibro.err.risk1.nrow = nrow(fibro.err.risk1.results)
fibro_crr_2_results = crr_cont_model_selection(fibro_dis, "fibro", risk=2, vars)
fibro_crr_risk2_results = crr_results(fibro_crr_2_results$fit_final, "FM", risk=2, fibro_crr_2_results$fit_final_vars)
fibro_crr_risk2_nrow = nrow(fibro_crr_risk2_results)
fibro_crr_3_results = crr_cont_model_selection(fibro_dis, "fibro", risk=3, vars)
fibro_crr_risk3_results = crr_results(fibro_crr_3_results$fit_final, "FM", risk=3, fibro_crr_3_results$fit_final_vars)
fibro_crr_risk3_nrow = nrow(fibro_crr_risk3_results)

## join all risk 1 results together in one data frame
crr_risk1_results_model_selection = rbind(dpnp_crr_risk1_results, fibro_crr_risk1_results, gad_crr_risk1_results, mdd_crr_risk1_results)
colnames(crr_risk1_results_model_selection) = c("Disease-State", "DC-Category", "Covariate", "Coefficient", "exp(coefficient)", "se(coefficient)", "Z-Statistic", "P-value", "Lower", "Upper")
crr_risk1_results_model_selection_xtable = subset(crr_risk1_results_model_selection, select=c("Disease-State", "Covariate", "exp(coefficient)", "Lower", "Upper"))

print(xtable(crr_risk1_results_model_selection_xtable, caption="CRR_Model_Selection_Results_for_DC_due_to_AE", label="tab:CRR_dcae_ms_results", align="|l|llccc|", hline.after=c(0, dpnp_crr_risk1_nrow, dpnp_crr_risk1_nrow+fibro_crr_risk1_nrow, dpnp_crr_risk1_nrow+fibro_crr_risk1_nrow+gad_crr_risk1_nrow, nrow(crr_risk1_results_model_selection_xtable))), include.rownames=F, include.colnames=F, caption.placement="top", add.to.row=list(po=1), command=c("\hline Disease-State & Risk-Factor & Subdistribution & HR & 95\%-CI\hline"), filepath=paste(path,"crr_dcae_results_model_selection.tex",sep=""))

## join all risk 2 results together in one data frame
crr_risk2_results_model_selection = rbind(dpnp_crr_risk2_results, fibro_crr_risk2_results, gad_crr_risk2_results, mdd_crr_risk2_results)
colnames(crr_risk2_results_model_selection) = c("Disease-State", "DC-Category", "Covariate", "Coefficient", "exp(coefficient)", "se(coefficient)", "Z-Statistic", "P-value", "Lower", "Upper")
crr_risk2_results_model_selection_xtable = subset(crr_risk2_results_model_selection, select=c("Disease-State", "Covariate", "exp(coefficient)", "Lower", "Upper"))
### join all risk 3 results together in one data frame

crr$_\text{risk3-results-model-selection} = \text{rbind}(\text{dpnp_crr_risk3-results}, \text{fibro_crr_risk3_results}, \text{gad_crr_risk3_results}, \text{mdd_crr_risk3-results})$

colnames(crr$_\text{risk3-results-model-selection}) = c("\text{Disease-State}", "\text{DC-Category}", "\text{Covariate}", "\text{Coefficient}", "\text{exp(coefficient)}", "\text{se(coefficient)}", "\text{Z-Statistic}", "\text{P-value}", "\text{Lower}", "\text{Upper}")

crr$_\text{risk3-results-model-selection-xtable} = \text{subset(crr$_\text{risk3-results-model-selection}, select=c("\text{Disease-State}", "\text{Covariate}", "\text{exp(coefficient)}", "\text{Lower}", "\text{Upper}")}$

### join all results together in one data frame

crr$_\text{all-results-model-selection} = \text{rbind}(\text{dpnp_crr_risk1-results, dpnp_crr_risk2-results, dpnp_crr_risk3-results, fibro_crr_risk1-results, fibro_crr_risk2-results, fibro_crr_risk3-results, gad_crr_risk1-results, gad_crr_risk2-results, gad_crr_risk3-results, gad_err_risk1_results, gad_crr_risk2_results, gad_err_risk3_results, mdd_crr_risk1-results, mdd_err_risk2_results, mdd_crr_risk3-results, mdd_crr_risk3-results)}$

colnames(crr$_\text{all-results-model-selection}) = c("\text{Disease-State}", "\text{DC-Category}", "\text{Covariate}", "\text{Coefficient}", "\text{exp(coefficient)}", "\text{se(coefficient)}", "\text{Z-Statistic}", "\text{P-value}", "\text{Lower}", "\text{Upper}")
A.9: Cox Regression Model Selection Program

coxph.cont.model.selection <- function(dataset, datatext, risk, var) {
  library(car)
  if (risk==1) {
    dataset$temp = recode(dataset$dscat, "2=0; -3=0")
  }
  else if (risk==2) {
    dataset$temp = recode(dataset$dscat_kly, "2=1; -3=0")
  }
  else if (risk==3) {
    dataset$temp = recode(dataset$dscat, "1=0; 2=0; 3=1")
  }
  dev <- pval <- varname <- index <- loglik.null <- loglik <- c(rep(0, length(var)))
  risk = c(rep(risk, length(var)))
  df = c(rep(1, length(var)))
  disease = c(rep(datatext, length(var)))

  for(j in 1:length(var)) {
    index[j] = j
    varname[j] = var[j]
    var_model = dataset[[varname[j]]]
    fit = coxph(Surv(visit-week, temp) ~ var_model, data=dataset)
    #fit = crr(dataset$visit.week, dataset$dscat_kly, dataset[, var[j]], failcode=risk)
    hr[j] = round(summary(fit, digits=4)$coef[, "exp(coef)"][,4]
loglik.null[j] = fitSloglik[1]

loglik[j] = fitSloglik[2]

dev[j] = round(-2*(loglik.null[j] - loglik[j]),4)

pval[j] = round(pchisq(dev[j], df[j], lower.tail=F),4)

results.step1 = data.frame(disease, risk, varname, index, hr, loglik.null, loglik, dev, pval)

varsig.step1 = as.vector(results.step1$varname[results.step1$pval<0.1])

varsig.step1.indexes = results.step1$index[results.step1$pval<0.1]

cat("Step_1_Analysis\n" , "Significant_variables_from_univariate_analysis\n")

print(varsig.step1)

print(varsig.step1.indexes)

cat("\n")

#cat(print(length(varsig.step1.indexes))

## Step 2: Fit full model with all sig vars from step 1 and try leaving out one variable at a time to see which vars are sig still

disease.step2 = rep(datatext, times=length(varsig.step1.indexes))

#cat(print(disease.step2)

#cat(" Risk is way too long of a vector so leave it out of data frame formulations

risk.step2 = rep(risk, times=length(varsig.step1.indexes))

#cat(print(risk.step2)

index.step2 = seq(along=varsig.step1.indexes)

dev <- pval <- varname <- index <- loglik <- hr <- c(rep(NA, length(varsig.step1.indexes)))

step1.full.formula <- as.formula(paste("Surv(visit.week,temp)\~",paste(varsig.step1, collapse="+")))

fit.full.step1 = coxph(step1.full.formula, data=dataset)

#cat("length of full model likelihood ",length(fit.full.step1$loglik), "\n")

loglik.step1 = fit.full.step1$loglik[2]

cat("log_likelihood_for_step_1_full_model_is_",loglik.step1, "\n")

loglik.step1.vec = rep(round(loglik.step1,4), times=length(varsig.step1.indexes))

#cat(print(loglik.step1.vec)

i = 1

for(j in 1:length(varsig.step1.indexes)) {

#cat("loop index = ",j,"\n")

varname[j] = varsig.step1[j]

varpos = which(var==varsig.step1[j])

varsig.now.indexes = varsig.step1.indexes[-j]
vars_model = var[varsig.new.indexes]
# print(vars_model)
new_formula = as.formula(paste("Surv(visit_week, temp)_~~", paste(vars_model, collapse="+")
))
# index[j] = j
fit = coxph(new_formula, data=dataset)
# shr[j] = round(summary.crr(fit, digits=4)$coef,"exp(coef)".4)
loglik[j] = fit$loglik[2]
# cat("loglik for loop index ",j," is " ,loglik[j],"\n")
# dev[j] = round(-2*(loglik_step1 - fit$loglik),4)
dev[j] = round(-2*(loglik_step1-loglik[j]),4)
# cat("dev for loop index ",j," is " ,dev[j],"\n")
pval[j] = pchisq(dev[j], df=1, lower.tail=F)
# cat("pval for loop index ",j," is " ,pval[j],"\n")

results_step2 = data.frame(varname, varsig_step1_indexes, loglik_step1_vec, loglik, dev, pval)
varsig_step2 = as.vector(results_step2$varname[results_step2$pval<0.1])
varsig_step2_indexes = results_step2$varsig_step1_indexes[results_step2$pval<0.1]
cat("Significant variables from step 2\n")
print(varsig_step2)
cat("\n")
# print(varsig_step2_indexes)
# cat("\n")

## Step 3: Try variables that were discarded in step 1 into full model from step 2

vars_to_try = vars[-varsig_step2_indexes]
vars_to_try_indexes = rep(NA, length(vars_to_try))
# vars_to_try_indexes = rep(NA, length(var)-length(varsig_step2_indexes))
dev <- pval <- varname <- loglik <- hr <- c(rep(NA, length(vars_to_try)))

step2_full_formula <- as.formula(paste("Surv(visit_week, temp)_~~", paste(vars_to_try, collapse="+")
))
fit_full_step2 = coxph(step2_full_formula, data=dataset)
loglik_step2 = fit_full_step2$loglik[2]
cat("log likelihood for step 2, full model is " ,loglik_step2, "\n")
loglik_step2_vec = rep(round(loglik_step2,4), times=length(vars_to_try))

for(j in 1:length(vars_to_try)) {
  # cat("loop index for step 3 = ",j,"\n")
  varname[j] = vars_to_try[j]
  vars_to_try_indexes[j] = which(var==vars_to_try[j])
vars_model = var[c(vars_to_try_indices[j], varsig_step2_indexes)]
new_formula = as.formula(paste("Surv(visit_week, temp) ~ ", paste(vars_model, collapse="+")
))
fit = coxph(new_formula, data=dataset)
# print(summary(fit))
loglik[j] = fit$loglik[2]
# cat("loglik for loop index ", j, " is ", loglik[j], 
# dev[j] = round(-2*(loglik_step1 - fit$loglik), 4)
# dev[j] = round(2*(loglik[j] - loglik_step2), 4)
# cat("dev for loop index ", j, " is ", dev[j], 
# pval[j] = pchisq(dev[j], df=1, lower.tail=F)
# cat("pval for loop index ", j, " is ", pval[j], 

results_step3 = data.frame(varname, vars_to_try_indices, loglik_step2_vec, loglik, dev, pval)

varsig_step3 = as.vector(results_step3$varname[results_step3$pval<0.1])
varsig_step3_indexes = results_step3$vars_to_try_indices[results_step3$pval<0.1]
cat("Significant variables from step 3
"
print(varsig_step3)
cat("\n")
# print(varsig_step3_indexes)
# cat("\n")

## Step 4: Check to make sure no term in the model can be omitted without significantly
## increasing the value of \(-2 \log L\)
step3_full_formula <- as.formula(paste("Surv(visit_week, temp) ~ ", paste(c(varsig_step2, varsig_step3), collapse="+")
))
fit_full_step3 = coxph(step3_full_formula, data=dataset)

loglik_step3 = fit_full_step3$loglik[2]
loglik_step3_vec = rep(round(loglik_step3, 4), times=length(c(varsig_step2_indexes, varsig_step3_indexes)))
dev <- pval <- varname <- index <- loglik <- shr <- c(rep(NA, length(c(varsig_step2_indexes, varsig_step3_indexes))))
varstochoose = c(varsig_step2, varsig_step3)
varstochoose_indexes = c(varsig_step2_indexes, varsig_step3_indexes)

for(j in 1:length(c(varsig_step2_indexes, varsig_step3_indexes))) {
  # cat("loop index for step 4 = ", j, 
  varname[j] = varstochoose[j]
  # varpos = which(var==varsig_step1[j])
  varsig_new_indexes = varstochoose_indexes[-j]}
vars_model = var[varsig.new.indexes]

# print(vars_model)

new_formula = as.formula(paste("Surv(visit.week, _temp) ~ ", paste(vars_model, collapse="+")))

# index[j] = j

fit = coxph(new_formula, data=dataset)

shr[j] = round(summary.crr(fit, digits=4)$coeff,"exp(coef)",4)

loglik[j] = fit$loglik[2]

# cat("loglik for loop index ", j, "is ", loglik[j], \\
# dev[j] = round(-2*(loglik_step1 - fit$loglik),4)

dev[j] = round(2*(loglik_step3 - loglik[j]),4)

# cat("dev for loop index ", j, "is ", dev[j], \\

pval[j] = pchisq(dev[j], df=1, lower.tail=F)

# cat("pval for loop index ", j, "is ", pval[j], \\

results_step4 = data.frame(varname, varstochoose_indexes, loglik_step3_vec, loglik, dev, pval)

varsig_step4 = as.vector(results_step4$varname[results_step4$pval < 0.1])

varsig_step4_indexes = results_step4$varstochoose_indexes[results_step4$pval < 0.1]

# cat("Significant variables from step 4")

print(varsig_step4)

# Step 5: Return data frame with parameters from final model

fit_final_formula <- as.formula(paste("Surv(visit.week, _temp) ~ ", paste(varsig_step4, collapse="+")))  

fit_final = coxph(fit_final_formula, data=dataset)

fit_final_vars = varsig_step4

fit_final_vars_indexes = varsig_step4_indexes

return(fit_final, fit_final_vars, fit_final_vars_indexes)

}

A.10: Cox Regression Model Selection Runner Program

source("R_funcs/coxph_cont_model_selection.R")

source("R_funcs/coxph_results.R")

path="/home/eric/Dropbox/diss_latex/ch4/tables/")

vars = c("ageyr", "pctchange_stand", "mxsevp_stand", "duration", "cau_ind", "country_cat.1", "sex_male", "trt_TRT")
## MOD

```r
mdd.coxph_1_results = coxph_cont_model_selection(mdd.dis, "mdd", risk=1, vars)
mdd.coxph.risk1_results = coxph_results(mdd.coxph_1_results$fit_final, "MDD", risk=1, mdd.coxph_1_results$fit_final_vars)
mdd.coxph.risk1_nrow = nrow(mdd.coxph.risk1_results)

mdd.coxph_2_results = coxph_cont_model_selection(mdd.dis, "mdd", risk=2, vars)
mdd.coxph.risk2_results = coxph_results(mdd.coxph_2_results$fit_final, "MDD", risk=2, mdd.coxph_2_results$fit_final_vars)
mdd.coxph.risk2_nrow = nrow(mdd.coxph.risk2_results)

mdd.coxph_3_results = coxph_cont_model_selection(mdd.dis, "mdd", risk=3, vars)
mdd.coxph.risk3_results = coxph_results(mdd.coxph_3_results$fit_final, "MDD", risk=3, mdd.coxph_3_results$fit_final_vars)
mdd.coxph.risk3_nrow = nrow(mdd.coxph.risk3_results)
```

## GAD

```r
gad.coxph_1_results = coxph_cont_model_selection(gad.dis, "gad", risk=1, vars)
gad.coxph.risk1_results = coxph_results(gad.coxph_1_results$fit_final, "GAD", risk=1, gad.coxph_1_results$fit_final_vars)
gad.coxph.risk1_nrow = nrow(gad.coxph.risk1_results)

gad.coxph_2_results = coxph_cont_model_selection(gad.dis, "gad", risk=2, vars)
gad.coxph.risk2_results = coxph_results(gad.coxph_2_results$fit_final, "GAD", risk=2, gad.coxph_2_results$fit_final_vars)
gad.coxph.risk2_nrow = nrow(gad.coxph.risk2_results)

gad.coxph_3_results = coxph_cont_model_selection(gad.dis, "gad", risk=3, vars)
gad.coxph.risk3_results = coxph_results(gad.coxph_3_results$fit_final, "GAD", risk=3, gad.coxph_3_results$fit_final_vars)
gad.coxph.risk3_nrow = nrow(gad.coxph.risk3_results)
```

## DPNP

```r
dpnp.coxph_1_results = coxph_cont_model_selection(dpnp.dis, "dpnp", risk=1, vars)
dpnp.coxph.risk1_results = coxph_results(dpnp.coxph_1_results$fit_final, "DPNP", risk=1, dpnp.coxph_1_results$fit_final_vars)
dpnp.coxph.risk1_nrow = nrow(dpnp.coxph.risk1_results)

dpnp.coxph_2_results = coxph_cont_model_selection(dpnp.dis, "dpnp", risk=2, vars)
dpnp.coxph.risk2_results = coxph_results(dpnp.coxph_2_results$fit_final, "DPNP", risk=2, dpnp.coxph_2_results$fit_final_vars)
dpnp.coxph.risk2_nrow = nrow(dpnp.coxph.risk2_results)
```
dpnp_coxph_3_results = coxph_cont.model_selection(dpnp_dis, "dpnp", risk=3, vars)
dpnp.coxph.risk3_results = coxph_results(dpnp.coxph_3_results$fit_final,"DPNP",risk=3, dpnp-
coxph.3_results$fit_final.vars)
dpnp.coxph.risk3_nrow = nrow(dpnp.coxph.risk3_results)

## FM
fibro_coxph_1_results = coxph_cont.model_selection(fibro_dis, "fibro", risk=1, vars)
fibro.coxph.risk1_results = coxph_results(fibro.coxph_1_results$fit_final,"FM",risk=1, fibro-
coxph.1_results$fit_final.vars)
fibro.coxph.risk1_nrow = nrow(fibro.coxph.risk1_results)

fibro.coxph.risk2_nrow = nrow(fibro.coxph.risk2_results)

fibro.coxph.risk3_nrow = nrow(fibro.coxph.risk3_results)

## join all risk 1 results together in one data frame
coxph.risk1_results.model.selection = rbind(dpnp.coxph.risk1_results, fibro.coxph.risk1-
results, gad.coxph.risk1_results, mdd.coxph.risk1_results)

colnames(coxph.risk1_results.model.selection) = c("Disease_State", "DC_Category", "Covariate",
"Coefficient", "exp(coefficient)", "se(coefficient)", "Z_Statistic", "P-value", "Lower",
"Upper")

coxph.risk1_results.model.selection_xtable = subset(coxph.risk1_results.model.selection,
select=c("Disease_State", "Covariate", "exp(coefficient)", "Lower", "Upper"))

print(xtable(coxph.risk1_results.model.selection_xtable, caption="CoxPH_Model_Selection-
Results_for_DC_dca_due_to_AE", label="tab:COXPH_dcae_ms_results", align="|l|l|c|c|c|", hline.
after=c(0,dpnp.coxph.risk1_nrow, dpnp.coxph.risk1_nrow+fibro.coxph.risk1_nrow,gad.coxph.
risk1_nrow+mdd.coxph.risk1_nrow, nrow(coxph.risk1_results.model.
selection_xtable)), include.rownames=F, include.colnames=F, caption.placement="top", add.
to=row=list(pos=list(-1), command="\\hline Disease_State & Risk_Factor & \multicolumn{2}{|c|}{95\% CI} & \multicolumn{1}{|c|}{hline}\film=pare(path,"coxph_dcae_results-
model_selection.tex",sep="\\))"
## join all risk 2 results together in one data frame

coxph_risk2_results_model.selection = rbind(dpnp.coxph.risk2_results, fibro.coxph.risk2_results, gad.coxph.risk2_results, mdd.coxph.risk2_results)

colnames(coxph_risk2_results_model.selection) = c("Disease-State", "DC-Category", "Covariate", "Coefficient", "exp(coefficient)", "se(coefficient)", "Z-Statistic", "P-value", "Lower", "Upper")

coxph_risk2_results_model.selection_xtable = subset(coxph_risk2_results_model.selection, select=c("Disease-State", "Covariate", "exp(coefficient)", "Lower", "Upper"))

print(xtable(coxph_risk2_results_model.selection_xtable, caption="Cox-PH Model Selection Results for DC due to LOE", label="tab:COXPH.dcloe.ms.results", align="|l|l|l|l|l|l|l|l|l|l|l|", hline.after=0, dpnp.coxph.risk2.nrow, dpnp.coxph.risk2.nrow+fibro.coxph.risk2.nrow, dpnp.coxph.risk2.nrow+fibro.coxph.risk2.nrow+gad.coxph.risk2.nrow, nrow(coxph_risk2_results_model.selection_xtable), include.rownames=F, include.colnames=F, caption.placement="top", add.to.rowlist(pos=list(1), command="\\hline Disease-State & Risk-Factor & Subdistribution HR & \multicolumn{2}{c|}{95\% CI} & \multicolumn{2}{c|}{95\% CI} & \multicolumn{2}{c|}{95\% CI} & \hline"), file=paste(path,"coxph_dcloe_results_model_selection.tex", sep=""))

## join all risk 3 results together in one data frame

coxph_risk3_results_model.selection = rbind(dpnp.coxph.risk3_results, fibro.coxph.risk3_results, gad.coxph.risk3_results, mdd.coxph.risk3_results)

colnames(coxph_risk3_results_model.selection) = c("Disease-State", "DC-Category", "Covariate", "Coefficient", "exp(coefficient)", "se(coefficient)", "Z-Statistic", "P-value", "Lower", "Upper")

coxph_risk3_results_model.selection_xtable = subset(coxph_risk3_results_model.selection, select=c("Disease-State", "Covariate", "exp(coefficient)", "Lower", "Upper"))

print(xtable(coxph_risk3_results_model.selection_xtable, caption="Cox-PH Model Selection Results for DC due to OTH", label="tab:COXPH.dcoth.ms.results", align="|l|l|l|l|l|l|l|l|l|l|l|", hline.after=0, dpnp.coxph.risk3.nrow, dpnp.coxph.risk3.nrow+fibro.coxph.risk3.nrow, dpnp.coxph.risk3.nrow+fibro.coxph.risk3.nrow+gad.coxph.risk3.nrow, nrow(coxph_risk3_results_model.selection_xtable), include.rownames=F, include.colnames=F, caption.placement="top", add.to.rowlist(pos=list(1), command="\\hline Disease-State & Risk-Factor & Subdistribution HR & \multicolumn{2}{c|}{95\% CI} & \multicolumn{2}{c|}{95\% CI} & \hline"), file=paste(path,"coxph_dcoth_results_model_selection.tex", sep=""))
## join all results together in one data frame

coxph_all_results_model_selection = rbind(dpnp_coxph_risk1_results, dpnp_coxph_risk2_results, dpnp_coxph_risk3_results, fibro_coxph_risk1_results, fibro_coxph_risk2_results, fibro_coxph_risk3_results, gad_coxph_risk1_results, gad_coxph_risk2_results, gad_coxph_risk3_results, mdd_coxph_risk1_results, mdd_coxph_risk2_results, mdd_coxph_risk3_results)

colnames(coxph_all_results_model_selection) = c("Disease-State", "DC-Category", "Covariate", "Coefficient", "exp(coefficient)", "se(coefficient)", "Z-Statistic", "P-value", "Lower", "Upper")

coxph_all_results_model_selection_xtable = subset(coxph_all_results_model_selection, select=c("Disease-State", "DC-Category", "Covariate", "exp(coefficient)", "Lower", "Upper")

print(xtable(coxph_all_results_model_selection_xtable, caption="Cox-PH_Model_Selection_Results", label="tab:COXPH_ms_results", align="||llccc|", hline.after=c(0,nrow(coxph_all_results_model_selection_xtable)), include.rownames=F, include.colnames=F, caption.placement="top", add.to.row=list(pos=list(-1),command=c("Disease_State","DC_Category","Risk_Factor","95\%_{CI}\multicolumn{2}{c|}{95\%_{CI}}\hline")), file=paste(path,"coxph_results_model_selection.tex",sep=""))