

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

5-2011

Marginal nonparametric inference for waiting times in multistage models : hypothesis testing and regression.

Douglas J. Lorenz
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

Recommended Citation

Lorenz, Douglas J., "Marginal nonparametric inference for waiting times in multistage models : hypothesis testing and regression." (2011). *Electronic Theses and Dissertations*. Paper 856.
<https://doi.org/10.18297/etd/856>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

MARGINAL NONPARAMETRIC INFERENCE FOR WAITING TIMES IN
MULTISTAGE MODELS: HYPOTHESIS TESTING AND REGRESSION

By

Douglas J. Lorenz
B.S., University of Louisville, 1999
M.A., University of Louisville, 2001
M.S.P.H. University of Louisville, 2002

A Dissertation
Submitted to the Faculty of the
Graduate School of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

Department of Mathematics
University of Louisville
Louisville, Kentucky

May 2011

MARGINAL NONPARAMETRIC INFERENCE FOR WAITING TIMES IN
MULTISTAGE MODELS: HYPOTHESIS TESTING AND REGRESSION

By

Douglas J. Lorenz
B.S., University of Louisville, 1999
M.A., University of Louisville, 2001
M.S.P.H. University of Louisville, 2002

A Dissertation Approved On

April 11, 2011

Date

by the following Dissertation Committee:

Somnath Datta

Co-Dissertation Director

Ryan Gill

Co-Dissertation Director

Guy Brock

Prasanna Sahoo

Kiseop Lee

DEDICATION

For my son, Gabriel Joseph Lorenz.

May he rest in eternal peace.

ACKNOWLEDGEMENTS

I thank my committee chairs, Dr. Somnath Datta and Dr. Ryan Gill, for their patience, timeless effort and invaluable insight throughout the course of my dissertation work. I also thank the members of my committee, Dr. Guy Brock, Dr. Prasanna Sahoo, and Dr. Kiseop Lee for additional insights and assistance.

I thank the Departments of Mathematics and Bioinformatics and Biostatistics and their members for the support and encouragement I received throughout my graduate study.

Finally, I thank my family, especially Melissa, Ophelia, Judah, and Zeke. Without the inspiration and support they have provided me, none of this would have been possible.

ABSTRACT

MARGINAL NONPARAMETRIC INFERENCE FOR WAITING TIMES IN MULTISTAGE MODELS: HYPOTHESIS TESTING AND REGRESSION

Douglas J. Lorenz

May 14, 2011

Marginal inference for waiting times in multi-stage time-to-event models is complicated by right censoring of observations as well as the prior history of events in the model. In general, complications arise due to the evolution of the censoring process in so called “calendar time”, contrasted with the evolution of the waiting time process conditional upon entry into a given stage. Developments in inference for survival data under dependent censoring have been extended to the multi-stage framework, and nonparametric estimators for the cumulative hazard function and survival function for waiting times analogous of the classical Nelson-Aalen and Kaplan-Meier estimators for survival data have been developed. These estimators were derived under the principle of weighting the basic at-risk and event counting processes by the inverse probability of censoring. We extend this concept to K -sample hypothesis testing and nonparametric regression, and define test statistics and regression coefficient estimators analogous to the log-rank test and Aalen’s nonparametric linear regression estimators for survival data. We examine the asymptotic distribution of these statistics, and justify their use via simulation studies and analyses of real data sets.

TABLE OF CONTENTS

	PAGE
DEDICATION	iii
ACKNOWLEDGEMENTS	iv
ABSTRACT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER	
I. INTRODUCTION	1
II. NOTATION AND PRELIMINARIES	8
Notation	9
A Model for the Censoring Hazard	10
Marginal Estimators of the Waiting Time Distribution	11
III. NONPARAMETRIC K -GROUP HYPOTHESIS TESTING	18
Expansion of Notation	19
Waiting Time Test Statistics for K Groups	21
Asymptotic Properties of the Test Statistic	23
Simulation Study – Validation	26
Simulation Study – Comparison	35
Analysis of Burn Patient Data	37
IV. NONPARAMETRIC REGRESSION	42
Estimators for Uncensored Data	43
IPCW Estimators	45
Simulation Study	47
Analysis of Bone Marrow Transplant Data	56

V. CONCLUSION	64
REFERENCES	70
APPENDIX	75
Proofs of Lemmas and Theorems	75
CURRICULUM VITAE	88

LIST OF TABLES

TABLE	PAGE
1. Censoring rates for test statistic simulations under the null hypothesis	28
2. Empirical size estimates of IPCW test and naive log-rank test at nominal 5% level. .	29
3. Description of stages in multi-stage burn data model.	37
4. Observed transitions of 154 burn patients	38
5. Disposition of 53 patients observed to enter stage 2 in the burn data model	39
6. Comparison of Kaplan-Meier and Satten-Datta stage 2 survival function estimates	40
7. Average bias for IPCW and naive coefficient estimators, simulation designs 1 and 2	51
8. Coverage probabilities for asymptotic 95% CI, simulation designs 1 and 2	54
9. Bias and coverage probabilities for asymptotic 95% CI, simulation design 3	55
10. Description of stages in multi-stage bone marrow transplant model.	57
11. Observed transitions of 137 bone marrow transplant patients	58

LIST OF FIGURES

FIGURE	PAGE
1. Empirical power of $\widehat{\mathbf{Z}}_1$ for 9 variants of simulation design 1	30
2. Empirical power of $\widehat{\mathbf{Z}}_1$ for 8 variants of simulation design 2	31
3. Power comparison of $\widehat{\mathbf{Z}}_1$ and \mathbf{Z}_1 for design 1 at $N = 500$	33
4. Power comparison of $\widehat{\mathbf{Z}}_1$ and \mathbf{Z}_1 for design 2 at $N = 500$	34
5. Power comparison of $\widehat{\mathbf{Z}}_1$ and the Fan-Datta test \bar{T}	36
6. Network of stages for multi-stage burn patient data.....	38
7. Stage 2 waiting times and survival functions for 4 groups of burn patients	40
8. IPCW coefficient estimators for selected variants of simulation designs 1-3.....	52
9. Naive coefficient estimators for selected variants of simulation designs 1-3.....	53
10. Network of stages for multi-stage bone marrow transplant data.....	58
11. Plots of stage 6 waiting times against covariates included in model (4.11).	60
12. Estimates of integrated regression coefficients from model (4.11) of bone marrow transplant data.....	61

CHAPTER I

INTRODUCTION

Multi-stage models provide a convenient framework for analyzing event history data, in which individuals proceed through a sequence of events in time. Multi-stage models generalize traditional survival or time-to-event data, in that individuals potentially experience a sequence of events rather than proceeding from a root stage to a single terminal absorbing stage. The events of multi-stage models are represented as stages and the occurrence of the events as transitions between the stages. Survival data, wherein individuals start at a root stage and progress to a single absorbing stage, provide the simplest example and are termed irreversible two-stage models in the nomenclature of multi-stage models. Other simple examples of multi-stage models include the competing risks model, in which individuals progress from a root stage to any of a set of absorbing stages. In biomedical applications, these states are generally referred to as an “alive” state (the root stage) and “death by cause X” (the set of absorbing stages). The three-stage illness-death model provides another simple example. Individuals progress from a root state (“well”) to an intermediary “illness” state and finally to an absorbing “dead” state. The multi-stage framework is suitably flexible to allow far more complex networks of stages. For example, the topology of the illness-death model can be made more complex by allowing direct transitions from the alive state to the dead state or by allowing transitions from the illness state back to the well state in an unlimited fashion.

The analysis of multi-stage data, also referred to as failure time data, is frequently complicated by right censoring, the partial observation of the progression through stages in the model caused by the exit of an experimental unit from observation before the full sequence of events has occurred. For simple survival data, traditional inferential procedures such as point and interval estimates of the cumulative hazard and survival functions, coefficients from parametric and semi-parametric regression models, and hypothesis test statistics, generally assume the statistical independence of survival times and censoring times. Practically speaking, this independence assumption provides reasonable assurance

that what an experimenter observes – the right-censored survival times – does not substantially differ from what actually occurs – the uncensored survival times – and that inference based on right-censored data will correspond with inference based on the uncensored data. When the independence of right-censoring and survival times cannot be safely assumed, inferential procedures requiring the independence assumption can be biased. For example, suppose that in an experiment censoring times and survival times are negatively correlated. Short survival times are then associated with long censoring times and are more likely to be fully observed. Conversely, long survival times are associated with short censoring times and more likely to be partially observed (censored). An estimator of the cumulative hazard function requiring the independence assumption would then overestimate the true cumulative hazard, as the fully observed survival times would tend to be shorter, indicating greater hazard, and the censored survival times would tend to be longer.

From an analytical perspective, there are many areas of interest for data derived from a multi-stage model – the estimation of stage-to-stage transition probabilities, hazard and survival functions for exit from a given stage, hazard and survival functions from a given stage to a particular subsequent stage, stage-occupation probabilities, etc. Presently, we will focus on the marginal analysis of waiting times for secondary, transient stages in a multi-stage model, defined as the time from entry into a given stage until stage exit and also referred to as sojourn or gap times. Waiting times are conceptually similar to survival times, in that one observes the time until a given event (stage exit) occurs, with the caveat that the time to stage exit is observed conditional upon entry into the given stage. Given this similarity, a naive approach to a marginal analysis of waiting times would be to apply survival data methods – the Nelson-Aalen estimator of the cumulative hazard function or Kaplan-Meier estimator of the survival function – on the set of individuals observed to enter a given stage, under the assumption that right censoring times are independent of stage transition times. However, such an independence assumption is insufficient to guarantee the applicability of methods for survival data to waiting times from multi-stage models. One must additionally assume the independence of stage transition times, the so-called semi-Markov property.

A heuristic argument can be used to demonstrate the additional need for the semi-Markov property. Let T and U denote the true entry and exit times for a given transient state in a multi-stage model, and let C denote the right-censoring time. Much of

the added complexity of analyzing multi-stage data stems from the fact that the waiting time process $U - T$ evolves conditional upon entry into the given stage (i.e. after random time T), whereas the censoring process evolves in “calendar time”, the time from which observation of the multi-stage model commences (i.e. “time 0”). In a marginal treatment of waiting times, we are in effect interested only in what happens after entry into the given stage, (i.e.) what happens after time T . Hence, the waiting time process, $U - T$, is censored not by C but by the censoring time after entry into the given stage, $C - T$. Assuming the independence of censoring and waiting times for a given stage in a marginal analysis thus corresponds to assuming the independence of $U - T$ and $C - T$. From this, it is easily seen that both independent censoring ($U, T \perp C$) and the semi-Markov property ($U \perp T$) are required. Hence, even when the waiting time $U - T$ can be assumed to be independent of the censoring time C in a multi-stage model, dependent censoring can be “induced” by dependence between U and T . This heuristic argument is not meant to imply that what happens before time T is completely irrelevant to the marginal analysis of waiting times; note that it is possible for individuals to be censored prior to entry into the stage of interest ($C - T < 0$). Survival data methods applied to waiting times from a multi-stage model would be necessarily carried out on the set of individuals observed to enter a given stage, completely ignoring individuals censored prior to stage entry that may have entered the given stage and posing an additional problem for their applicability.

In practice, the semi-Markov property can be restrictive and implausible for multi-stage models and in the marginal analysis of waiting times. Inference for survival data under dependent censoring and waiting times from multi-stage models under dependent censoring and/or transition times has generally focused on adjusting for dependent censoring in some capacity. A method that has received particular focus is known as inverse probability of censoring weighting (IPCW), first introduced by Robins and Rotnitzky [39] in the analysis of survival data. To describe IPCW, we first note that many statistics for survival data are functions of two counting processes evolving in time, (1) the event counting process, which counts the number of events of interest (failures) that have occurred up to a given time, and (2) the at-risk counting process, which counts the number of experimental units at risk for the event of interest at a given time. (The counting process formulation for the analysis of survival data was first proposed by Aalen [1] in a landmark paper, and has received considerable attention since.) The IPCW method operates by weighting these basic counting process by the inverse probability of censoring, thereby

weighting the statistics that are functions of these counting processes. Thus, observations more/less likely to be censored are given greater/lower weight, in turn adjusting for the potential bias introduced by dependent censoring. The IPCW method has root in the Horvitz-Thompson estimator [19], a popular method in the analysis of sample surveys, which weights observations by the inverse probability of an observation being selected.

The IPCW method has been used to develop dependent-censoring analogues of several regular estimators for survival data. Robins and Rotnitzky [40] developed coefficient estimators for semiparametric regression models under dependent censoring analogous to well-known estimators from the Cox proportional hazards model [9]. Satten and Datta [41] demonstrated that the Kaplan-Meier estimator [24] of the survival function had representation as an IPCW average, and Satten, Datta and Robins [42, 43] developed estimators of the cumulative hazard and survival functions under dependent censoring as analogues of the classical Nelson-Aalen [1, 36] and Kaplan-Meier estimators. Robins and Finkelstein [38] adapted IPCW semiparametric regression estimators to develop IPCW Kaplan-Meier estimators and score tests for two group comparisons. We note that in this paper, IPCW was used to correct not only for dependent censoring but also non-compliance with a given therapeutic regimen in a two-arm study of AIDS treatment. Cain and Cole [7] extended these results to correct time-varying non-compliance with treatment. Howe, et al [21] have noted limitations in the estimation of survival via IPCW in the presence of strong selection bias. Recently, Datta, et al [10] developed a broad theory for IPCW U -statistics with applications to testing problems for right-censored data. We note that weighted estimation for survival data was first considered by Koul, Susarla, and Van Ryzin [28], who developed weighted OLS regression estimators for randomly right censored survival data.

Weighting by the inverse probability of censoring has also been used in the analysis of multi-stage data. The IPCW approach has previously been applied to multi-stage data to develop estimators of transition hazards and stage occupation probabilities [11, 13] and to prove the validity of an Aalen-Johansen-type estimator [5] of the integrated transition hazard matrix for non-Markov models [12]. Gunnes, Borgan, and Aalen [18] conducted a detailed examination of the Aalen-Johansen and Datta-Satten estimators of stage occupation probabilities, and noted that the Datta-Satten estimator remained approximately unbiased while the Aalen-Johansen estimator under- or over-estimated stage occupation probabilities based on dependent censoring patterns. The authors did note that in their simulations, differences between the Datta-Satten and Aalen-Johansen estimators

were of practical importance only under very selective censoring mechanisms. For the analysis of waiting times from a multi-stage model, Matsuyama and Yamaguchi [35] introduced IPCW estimators of marginal survival times for dependent competing risks. These estimators were a special case of survival function estimators for waiting times from general multi-stage models, which were developed by Satten and Datta [44] as IPCW analogues of the Nelson-Aalen and Kaplan-Meier estimators for survival data. The application of IPCW methods to multi-stage data have not been limited to being marginal in scope. Lahkal-Chaeib, Cook and Lin [29] examined the joint distribution of waiting times from multi-stage models, and in particular developed nonparametric estimators for correlations between waiting times from consecutive states based on Kendall's τ .

We briefly note that the IPCW method does not represent the only method for adjusting for dependent censoring for survival data or dependent transition times in a multi-stage model, nor the only method for reweighting to correct for bias. Wang and Wells [47] and Lin, Sun, and Ying [32] introduced marginal estimators for waiting times in the transient (i.e. second) stage in a three-stage illness death model. Huang [22] and Strawderman [46] separately developed accelerated waiting time models for multi-stage data as natural extensions of the accelerated failure time model for survival data. Schaubel and Cai [45] proposed regression models for the waiting time hazard function (termed gap time hazard by the authors). In each of these papers, the authors arrived at model estimators through appropriate reweightings of regular estimating equations for survival data.

In what follows, we employ the principle of IPCW to develop K -sample test statistics and nonparametric regression estimators for waiting times from a multi-stage model. The test statistics we introduce serve as analogues of the Fleming-Harrington [16] class of K -group tests for survival data, which includes the popular log-rank test [34]. Hypothesis testing under dependent censoring for survival data has previously been considered by Lin, Robins, and Wei [31], who introduced a two-sample test statistic by considering a bivariate location-shift model with unspecified underlying distribution for the failure and censoring times. Lin and Ying [33] proposed weighted log-rank and Pepe-Fleming [37] type statistics for waiting times involving reweighting and conditioning on threshold values for stage entry times. Fan and Datta [15] developed IPCW Mann-Whitney statistics for two-group comparisons of waiting times from multi-stage models. Additionally, the asymptotic theory for the regression models mentioned in the previous paragraphs can provide K -sample hypothesis test statistics when the model covariate is K -categorical. As noted above,

Robins and Finkelstein [38], in developing IPCW semi-parametric regression and Kaplan-Meier estimators, additionally proposed IPCW tests for survival data as score tests derived from the estimation of the regression coefficients.

The nonparametric regression estimators for waiting times from a multi-stage model we propose are analogues of the nonparametric regression estimators for survival data introduced by Aalen [2, 3, 4]. This model has received considerably less attention than other regression models for survival data, such as the semi-parametric Cox model [9], fully parametric regression models, and the accelerated failure time (AFT) model. There are notable reasons for this lack of attention. Aalen’s model directly models the hazard function, rather than an exponential or logarithmic transform of the hazard function as with the Cox and AFT models. While this makes Aalen’s model conceptually simpler to define, negative estimates of the hazard are possible. Regression matrices of less than full column rank can cause erratic behavior of hazard estimates provided by Aalen’s model, particularly at later times of observation when the risk set is small. Nevertheless, Aalen’s nonparametric linear model provides unique flexibility in modeling the hazard as it not only permits time-varying covariates (like the Cox and AFT models) but defines its regression coefficients as functions of time rather than as static values, (i.e.) the regression coefficients are time-varying as well. This unique flexibility uncovers an additional drawback in that Aalen’s model estimates integrated regression coefficients which can be difficult to interpret. However, kernel smoothing techniques can be used to generate smoothed estimates of the regression coefficient functions and ad hoc solutions to the other difficulties – setting a lower bound of zero for model-generated hazard estimates and skipping failure times for which the regression matrix is rank deficient [20] – generally perform well and mitigate these problems to some degree. Other than the development of tests of covariate effects from Aalen’s model [17, 30], limited further research into this model has been conducted. No direct extensions of Aalen’s model to multi-stage data have been proposed.

We detail our IPCW test statistics and nonparametric regression estimators for waiting times from multi-stage models in the following chapters, which are organized as follows. In Chapter II, we develop our notation, review the principle of IPCW and the waiting time cumulative hazard and survival function estimators of Satten and Datta [44], define a model for estimating the censoring hazard which is necessary for the practical use of our proposed statistics, and establish technical results for use in subsequent chapters. In Chapter III, we introduce test statistics for the comparison of waiting time distributions

among K -groups. In Chapter IV, we review Aalen's nonparametric linear model for survival data and develop nonparametric regression estimators for waiting times. Both Chapters III and IV contain the results of simulation studies exploring the asymptotic properties of our test statistics and regression estimators as well as the analysis of real data sets demonstrating their use. The dissertation concludes with a discussion of our results and suggestions for further avenues of research in Chapter V.

CHAPTER II

NOTATION AND PRELIMINARIES

In this chapter we introduce notation for multi-stage models and review the estimators of Satten and Datta [44] for the cumulative hazard and survival function for waiting times in a multi-stage model. In so doing, we review the principle of weighting the basic counting processes by the inverse probability of censoring, first introduced by Robins and Rotnitzky [39, 40]. The purpose of this review is to provide motivation for our proposed test statistics and nonparametric regression estimators for waiting times, and to establish technical results upon which results later in the text will be based. We note that some of the proofs in this chapter follow directly from corresponding proofs previously laid out for survival data [12, 41, 43], but have yet to be formally established for waiting times in a multi-stage model. Much like the cumulative hazard and survival function estimators, the test statistics and regression estimators we propose are simple analogues of statistics commonly used for survival data, with basic counting processes for uncensored data replaced by their IPCW equivalents. The discussion below largely follows that of Datta and Satten [11], Satten, Datta, and Robins [42, 43], and Satten and Datta [44]. Specifically, we restrict our attention to continuous waiting time distributions and acyclic network structures, (i.e.) networks with no cycles among subsets of the stages. As discussed by Satten and Datta [44], a cyclic network can be exploded into an acyclic network by considering the unique times of entry into a given stage, (e.g.) the first, second, third, etc. entries into a given stage of a cyclic network can be reconstructed as unique stages in the model. The authors further note that an exploded stage of an acyclically restructured network subsequently can be pooled to generate hazard and survival function estimates for “general” stage exit (as opposed to order-specific estimates), so there is no loss of generality.

We briefly note that the notation and preliminary results outlined in this chapter are in the context of a single sample. This notation will be directly applicable for the nonparametric regression coefficient estimators proposed in Chapter IV. Chapter III will deal with hypothesis testing in the K -sample setting, and will necessarily require extended

notation that will be discussed in that chapter.

Notation

Since we will consider potentially right-censored data from multi-stage models, it will be necessary in defining our notation to delineate “true” quantities from their censored counterparts, (i.e.) to differentiate what actually occurs from what is observed by the experimenter/analyst. In general, we will denote true quantities and functions thereof with an asterisk (*) and their observed (censored) equivalents will not be annotated.

Let i ($1 \leq i \leq n$) index individuals and j ($1 \leq j \leq J$) index the stages in the model. Let T_{ij}^* denote the time individual i enters stage j , and let U_{ij}^* denote the corresponding exit time. Note that $T_{ij}^* = \infty$ if stage j is never entered and $U_{ij}^* = \infty$ if stage j is never entered or never exited, for example, if stage j is absorbing. Define $T_i^* = \max_j \{T_{ij}^* \mid T_{ij}^* < \infty\}$ as the time of the last transition for individual i . Let $X_{ij} = I[T_{ij}^* < \infty]$ be the indicator of individual i ever entering stage j . The values T_{ij}^* , U_{ij}^* , T_i^* , and X_{ij} can go unobserved due to right censoring. Let C_i denote the censoring time for individual i . Define $\gamma_{ij} = I[C_i \geq T_{ij}^*]$ to be the indicator of individual i having been observed to enter stage j , and $\delta_{ij} = I[C_i \geq U_{ij}^*]$ the indicator of individual i having been observed to leave stage j . Let $T_i = \min(T_i^*, C_i)$ be the time of the final observed transition. Define $T_{ij} = \min(T_{ij}^*, C_i)$ if the data at time t do not imply that $X_{ij} = 0$ and ∞ otherwise. Note that $T_{ij} = \infty$ if, for example, an individual proceeds down a path in the network from which stage j is unreachable. Let $U_{ij} = \min(U_{ij}^*, C_i)$ if j is not absorbing and the data at time t do not imply that $X_{ij} = 0$ and ∞ otherwise. For the moment, we assume that the observed data consist of i.i.d. replicates of $\{T_{ij}, U_{ij}, \gamma_{ij}, \delta_{ij}, 1 \leq j \leq J\}$.

In addition to the entry and exit times and indicators, we observe a set of possibly time-varying covariates that may impact both the censoring and transition hazards. Denote the p -vector of these “external” covariates for individual i by $\mathbf{X}_i^1(t)$, and the history of these values up to and including time t by $\overline{\mathbf{X}}_i^1(t)$. Let $\mathcal{F}_t = \{T_{ij}^* \mid T_{ij}^* \leq t\}$ denote the history of transitions up to and including time t . Define the vector of “internal” covariates as $\mathbf{X}_i^2(t)$, which are functions only of \mathcal{F}_t , and denote the history of these covariates as $\overline{\mathbf{X}}_i^2(t)$. The collection of internal and external covariates for individual i is denoted by $\mathbf{X}_i(t)$ and its corresponding history as $\overline{\mathbf{X}}_i(t)$.

In the present chapter, we will review the Satten and Datta [44] estimators of the

waiting time cumulative hazard and survival functions, for which we provide definitions. Let $S_j(t) = P[U_{ij}^* - T_{ij}^* > t | X_{ij} = 1]$ be the survival function for stage j waiting times and let $\Lambda_j(t) = \int_0^t \lambda_j(s) ds$ be the corresponding cumulative hazard function, where $\lambda_j(t) dt = P[U_{ij}^* - T_{ij}^* \in [t, t + dt] | U_{ij}^* - T_{ij}^* \geq t]$ denotes the hazard rate function, the conditional probability of stage j exit over an infinitesimal interval given that stage j exit has not occurred prior to t units of time after stage entry.

A Model for the Censoring Hazard

The principle behind estimation by IPCW for survival data introduced by Robins and Rotnitzky [39, 40] is fairly simple. Many traditional estimators for survival data are functions of two counting processes – the event process, which counts the number of events (i.e. deaths or failures) that occur, and the at-risk process, which counts the number of individuals under consideration that remain at risk of the event. The IPCW estimator for a given population quantity, such as the survival function, is achieved by replacing the counting processes composing a traditional estimator for uncensored data with weighted equivalents, where the weights are defined by some estimate of the inverse probability of censoring. As will be seen, the IPCW estimators have useful statistical properties but remain applicable when independent censoring cannot be assumed.

In order to formally define IPCW estimators, we must develop a conceptual model for the censoring hazard. Following Robins and Rotnitzky [39, 40] and Satten, Datta, and Robins [42, 43], we assume a censoring mechanism in our multi-stage model in which knowledge of future transition times does not affect the censoring hazard. Specifically, letting $\lambda^c(t | \cdot)$ denote the censoring hazard, we assume that

$$\lambda^c(t | \bar{\mathbf{X}}_i^1(t-), \mathcal{F}_\infty) = \lambda^c(t | \bar{\mathbf{X}}_i^1(t-), \mathcal{F}_{t-}),$$

where \mathcal{F}_t is as defined previously. Note that this assumption indicates that the censoring hazard, as a function of the history of the external covariates ($\bar{\mathbf{X}}_i^1(t-)$) and the history of transitions (\mathcal{F}_t), is impacted only by the history of transition times up to just before time t , (\mathcal{F}_{t-}). We further assume that the set of internal covariates, $\mathbf{X}_i^2(t)$, which are functions only of \mathcal{F}_t , are chosen such that

$$\lambda^c(t | \bar{\mathbf{X}}_i^1(t-), \mathcal{F}_{t-}) = \lambda^c(t | \bar{\mathbf{X}}_i^1(t-), \bar{\mathbf{X}}_i^2(t-)).$$

Specifically, we are asserting that the history of the internal covariates $(\bar{X}_i^2(t-))$ fully accounts for any impact that the history of transition times (\mathcal{F}_{t-}) may have on the censoring hazard. In other words, any information contained in \mathcal{F}_{t-} not contained in $\bar{X}_i^2(t-)$ does not impact the censoring hazard. Define the cumulative hazard of censoring in the usual fashion, as $\Lambda^c(t | \bar{X}_i(t)) = \int_0^t \lambda^c(s | \bar{X}_i(s-)) ds$ and the product integral of the censoring hazard as $K_i(t) = \prod_{s \leq t} [1 - d\Lambda^c(s | \bar{X}_i(s))]$. In what follows, we will typically write $\lambda_i^c(t)$ to denote $\lambda^c(t | \bar{X}_i(t-))$. A functional form for $\lambda_i^c(t)$ will be developed in the following section. For now, it is sufficient to stipulate that the censoring hazard at time t is defined as a function of the covariates (internal and external) up to just before t .

Marginal Estimators of the Waiting Time Distribution

For uncensored data, in which the stage entry and exit times are fully observed, one could estimate $S_j(t)$ and $\Lambda_j(t)$ using the Nelson-Aalen and Kaplan-Meier estimators for waiting times. To this end, define the uncensored data counting process for exits from stage j as $N_j^*(t) = \sum_{i=1}^n I[U_{ij}^* - T_{ij}^* \leq t, X_{ij} = 1]$, and define the corresponding at risk process as $Y_j^*(t) = \sum_{i=1}^n I[U_{ij}^* - T_{ij}^* \geq t, X_{ij} = 1]$. The uncensored data estimators for the stage j waiting time cumulative hazard and survival functions are then

$$\Lambda_j^*(t) = \int_0^t \frac{dN_j^*(s)}{Y_j^*(s)}, \quad S_j^*(t) = \prod_{s \leq t} \left(1 - \frac{dN_j^*(s)}{Y_j^*(s)} \right), \quad (2.1)$$

where Π denotes the product integral. The estimators in (2.1) cannot be used for censored multi-stage data, but establish a useful foundation upon which to base IPCW estimators. Estimators for right-censored data could be derived by replacing in (2.1) the uncensored data counting processes with their censored data equivalents:

$N_j(t) = \sum_{i=1}^n I[U_{ij} - T_{ij} \leq t, \delta_{ij} = 1]$ and $Y_j(t) = \sum_{i=1}^n I[U_{ij} - T_{ij} \geq t, \gamma_{ij} = 1]$. However, as noted in Chapter I, the applicability of such estimators generally requires the assumption of independent censoring and the semi-Markov property.

As noted in the previous section, the IPCW estimators of the cumulative hazard and survival function for waiting times are derived by replacing the traditional counting processes (censored or uncensored) with their weighted equivalents. The $K_i(t)$ defined in the previous section serve as our probabilities of censoring, and thus our inverse probability

of censoring weights. Define the weighted event and at-risk counting processes as follows:

$$\begin{aligned}\bar{N}_j(t) &= \sum_{i=1}^n \frac{I[U_{ij} - T_{ij} \leq t, \delta_{ij} = 1]}{K_i(U_{ij}-)} \\ \bar{Y}_j(t) &= \sum_{i=1}^n \frac{I[U_{ij} - T_{ij} \geq t, \gamma_{ij} = 1]}{K_i(T_{ij} + t-)}.\end{aligned}$$

Note that we have weighted the censored data counting processes $N_j(t)$ and $Y_j(t)$, since weighting the uncensored data counting processes would still require uncensored data and be of little practical use. However, the justification for the use of \bar{N}_j and \bar{Y}_j lies in a correspondence with the uncensored data counting processes N_j^* and Y_j^* defined by the following theorem.

Theorem II.1 Let t be such that $E[K_i^{-2}(t)] < \infty$ for $1 \leq i \leq n$. Then

$$\begin{aligned}E[\bar{N}_j(t)] &= E[N_j^*(t)], \\ E[\bar{Y}_j(t)] &= E[Y_j^*(t)].\end{aligned}$$

Before proving Theorem II.1, we prove a useful technical lemma, that will be used in proofs of future theorems as well. The proof of the lemma involves an application of the Duhamel equation (cf. Andersen, et al [6], Thm. II.6.2), and follows the arguments laid out in a similar proof by Satten, Datta, and Robins [43] for survival data (as does the proof of Theorem II.1). Before proceeding, define $N_i^c(t) = I[C_i \leq t, C_i \leq T_i^*]$ to be the counting process of censoring events and $M_i^c(t) = N_i^c(t) - I[T_i \geq t]\Lambda_i^c(t)$ to be the associated martingale, where $\Lambda_i^c(t)$ is the cumulative hazard of censoring as previously defined. Note that in defining the censoring event counting process and martingale we have used t to denote time, as with the counting processes associated with stage exit. Despite this, it is important to distinguish that the stage exit process evolves conditional upon stage entry (i.e. after time T_{ij} for individual i) whereas the censoring process evolves in so-called “calendar time”, after observation begins in the multi-stage system. In other words, the t in $\bar{N}_j(t)$ refers to time t after entry into stage j , whereas the t in $N_i^c(t)$ refers to time t after observation of the multi-stage system begins (at time 0). We will carefully distinguish these two time scales (calendar time and time after stage entry) as needed when we lay out our results. We then have the following:

Lemma II.2 For $t \leq T_i^*$,

$$\frac{I[C_i \geq t]}{K_i(t-)} = 1 - \int_0^{t-} \frac{I[C_i \geq s]}{K_i(s)} dM_i^c(s).$$

Proof. Note that $I[C_i \geq t] = \prod_{s < t} [1 + dX_i(s)]$, where $X_i(s) = -I[C_i \leq s]$. For $s < t \leq T_i^*$, we have $X_i(s) = -I[C_i \leq s, C_i \leq T_i^*] = -N_i^c(s)$. On the set $\{C_i \geq T_i^*\}$, we have $K_i(t-) = \prod_{s < t} [1 - dX_i'(s)]$, where $X_i'(s) = I[T_i \geq s] \Lambda_i^c(s)$. By the Duhamel equation, we get

$$\frac{I[C_i \geq t]}{K_i(t-)} = 1 + \int_0^{t-} \frac{I[C_i \geq s]}{K_i(s)} (dX_i(s) + dX_i'(s)).$$

Note that $X_i(s) + X_i'(s) = -N_i^c(s) + I[T_i \geq s] \Lambda_i^c(s) = -M_i^c(s)$. Hence,

$$\frac{I[C_i \geq t]}{K_i(t-)} = 1 - \int_0^{t-} \frac{I[C_i \geq s]}{K_i(s)} d\widehat{M}_i^c(s).$$

■

We make a brief note that this proof also carries through if we replace $K_i(t-)$ by some estimator $\widehat{K}_i(t-)$, with the modification that the right hand side integral is taken with respect to $d\widehat{M}_i^c(s)$, where $\widehat{M}_i^c(s) = N_i^c(s) - I[T_i \geq s] \widehat{\Lambda}_i^c(s)$ and $\widehat{\Lambda}_i^c(t)$ is the estimator of $\Lambda_i^c(t)$ generating $\widehat{K}_i(t)$ via the product integral. The quantity $\widehat{M}_i^c(s)$ also does not define a martingale, but will be useful in later results. Note also that the proof of Lemma II.2 made no assumption of the relationship between the censoring times C_i and the transition times T_{ij} , (i.e.) neither the assumption of independent censoring nor the semi-Markov property were required. Lemma II.2 will be used repeatedly in several later proofs.

Proof of Theorem II.1. To prove the first equality, note that U_{ij}^* satisfies the conditions of Lemma II.2, i.e. $U_{ij}^* \leq T_i^*$. Hence,

$$\frac{I[C_i \geq U_{ij}^*]}{K_i(U_{ij}^*-)} = 1 - \int_0^{U_{ij}^*-} \frac{I[C_i \geq s]}{K_i(s)} dM_i^c(s).$$

Multiplying both sides by $I[U_{ij}^* - T_{ij}^* \leq t]$ produces

$$\frac{I[U_{ij}^* - T_{ij}^* \leq t, C_i \geq U_{ij}^*]}{K_i(U_{ij}^*-)} = I[U_{ij}^* - T_{ij}^* \leq t] - \int_0^{U_{ij}^*-} \frac{I[U_{ij}^* - T_{ij}^* \leq t]}{K_i(s)} dM_i^c(s),$$

where the $I[C_i \geq s]$ term in the integrand on the right hand side can be removed since the martingale $M_i^c(s)$ vanishes for $C_i < s$. On the left hand side, the indicator function in the

numerator is equivalent to $I[U_{ij} - T_{ij} \leq t, C_i \geq U_{ij}^*]$. Further, on the set $\{C_i \geq U_{ij}^*\}$, we have that $K_i(U_{ij}^*-) = K_i(U_{ij}-)$ and trivially on the right hand side, $I[U_{ij}^* - T_{ij}^* \leq t] = I[U_{ij}^* - T_{ij}^* \leq t, X_{ij} = 1]$. Hence, we have

$$\bar{N}_{ij}(t) = N_{ij}^*(t) - \int_0^{U_{ij}^*-} \frac{I[U_{ij}^* - T_{ij}^* \leq t]}{K_i(s)} dM_i^c(s).$$

We note that $I[U_{ij}^* - T_{ij}^* \leq t]/K_i(s)$ is \mathcal{F}_s -predictable, and hence the integral on the right hand side is a zero-mean martingale (since $M_i^c(s)$ is). The first equality follows by taking the expectation of both sides, under the stated moment condition on $K_i(s)$.

The proof of the second equality follows similarly, by noting that on the set $\{U_{ij}^* - T_{ij}^* \geq t\}$, $T_{ij}^* + t$ satisfies for the conditions of Lemma I.1 so that

$$\frac{I[C_i \geq T_{ij}^* + t]}{K_i(T_{ij}^* + t-)} = 1 - \int_0^{T_{ij}^*+t-} \frac{I[C_i \geq s]}{K_i(s)} dM_i^c(s).$$

(For times after U_{ij}^* , both $Y_{ij}^*(t)$ and $\bar{Y}_{ij}(t)$ are zero.) Multiplying both sides by $I[U_{ij}^* - T_{ij}^* \geq t]$ produces

$$\frac{I[U_{ij}^* - T_{ij}^* \geq t, C_i \geq T_{ij}^* + t]}{K_i(T_{ij}^* + t-)} = I[U_{ij}^* - T_{ij}^* \geq t] - \int_0^{T_{ij}^*+t-} \frac{I[U_{ij}^* - T_{ij}^* \geq t]}{K_i(s)} dM_i^c(s).$$

The numerator on the left hand side is equivalent to $I[U_{ij} - T_{ij} \geq t, C_i \geq T_{ij}^*]$, and on the set $\{C_i \geq T_{ij}^*\}$, we have that $K_i(T_{ij}^* + t-) = K_i(T_{ij} + t-)$ and we trivially have that $I[U_{ij}^* - T_{ij}^* \geq t] = I[U_{ij}^* - T_{ij}^* \geq t, X_{ij} = 1]$. Hence, we have that

$$\bar{Y}_{ij}(t) = Y_{ij}^*(t) - \int_0^{T_{ij}^*+t} \frac{I[U_{ij}^* - T_{ij}^* \geq t]}{K_i(s)} dM_i^c(s).$$

Again, the result follows by noting that $I[U_{ij}^* - T_{ij}^* \geq t]/K_i(s)$ is \mathcal{F}_s -predictable, the right hand side integral is a zero mean martingale, and taking expectations. ■

The large sample equivalence of the uncensored data and IPCW counting processes indicates that inference based on the IPCW processes will be asymptotically equivalent to inference based on the uncensored data processes, a clearly desirable property. In essence, \bar{N}_j and \bar{Y}_j “estimate” the uncensored data counting processes N_j^* and Y_j^* , which are unobservable when right-censoring is present. However, the reweighted counting processes

\bar{N}_j and \bar{Y}_j are of little practical use since they are based upon the quantity $K_i(t)$ which is typically unknown. The application of the IPCW counting processes requires an estimator for $K_i(t)$. In introducing IPCW, Robins and Rotnitzky [39, 40] used the well-known Cox proportional hazards model [9]. Datta and Satten [12] suggested the use of Aalen's linear model [2, 3, 4] for the added flexibility of time-varying model coefficients, and we adopt its usage here. We briefly note that in suggesting the use of Aalen's linear model, Datta and Satten argue that the potential negative side effects of its use (regression matrices with less than full column rank and negative estimates of the hazard function) do not impact the estimation of the cumulative hazard or survival function for stage waiting times [12, 43, 44].

We now define Aalen's linear model for the censoring hazard in our current framework: $\lambda_i^c(t) = \beta_0^c(t) + \sum_{m=1}^p \beta_m^c(t) X_{im}(t)$, where the $\beta_m^c(t)$ are the coefficient functions and m indexes both internal and external covariates. Note, in particular, that both the regression coefficients $\beta_m^c(t)$ and covariates $X_{im}(t)$ are time-varying functions rather than static values, the source of the unique flexibility of Aalen's model. Define the integrated regression coefficients $B_m^c(t) = \int_0^t \beta_m^c(s) ds, 0 \leq m \leq p$. Aalen's linear model estimates the functions $\mathbf{B}^c(t) = (B_0^c(t), \dots, B_p^c(t))$ by

$$\hat{\mathbf{B}}^c(t) = \sum_{i=1}^n I(T_i \leq t)(1 - \delta_i) \mathbf{A}^{-1}(T_i) \mathbf{X}_i(T_i),$$

where $\delta_i = I[T_i^* < C_i]$ and $\mathbf{A}(t) = \sum_{i=1}^n I(T_i \geq t) \mathbf{X}_i(t) \mathbf{X}_i^T(t)$. Note that we have defined the estimator $\hat{\mathbf{B}}^c(t)$ in terms of a specific generalized inverse for $\mathbf{A}(t)$, namely $\mathbf{A}^{-1}(t) \mathbf{X}_i(t)$ which arises from the least squares principle, and that any generalized inverse of $\mathbf{A}(t)$ will suffice. We are now able to estimate the censoring hazard and our inverse probability of censoring weights, $\hat{\Lambda}_i^c(t) = \int_0^t \mathbf{X}_i^T(s) d\hat{\mathbf{B}}^c(s)$ and $\hat{K}_i(t) = \prod_{s \leq t} (1 - d\hat{\Lambda}_i^c(s))$.

We now define the counting processes $\hat{N}_j(t) = I[U_{ij} - T_{ij} \leq t, \delta_{ij} = 1] / \hat{K}_{ij}(t)$ and $\hat{Y}_j(t) = I[U_{ij} - T_{ij} \geq t, \gamma_{ij} = 1] / \hat{K}_{ij}(t)$, which are completely functions of the observable data. The Satten and Datta [44] estimators for the cumulative hazard and survival functions for stage j waiting times are then

$$\hat{\Lambda}_j(t) = \int_0^t \frac{d\hat{N}_j(s)}{\hat{Y}_j(s)}, \quad \hat{S}_j(t) = \prod_{s \leq t} \left(1 - \frac{d\hat{N}_j(s)}{\hat{Y}_j(s)} \right). \quad (2.2)$$

As indicated above, the IPCW estimators are simple analogues of the uncensored data estimators in (2.1), with the uncensored data counting processes replaced by their IPCW

equivalents. These estimators, as functions of the IPCW counting processes, are calculable of the observable data. The consistency of these estimators for $\Lambda_j(t)$ and $S_j(t)$ is established in the following theorem.

Theorem II.3 *Let τ be such that $\tau > T_i$ for all i . Let t be such that $E[K_i^{-2}(t)] < \infty$ for $1 \leq i \leq n$ and assume that for all $T < \tau$, $\sup_{t \leq T} n^{-1} \sum_{i=1}^n |\widehat{K}_i(t) - K_i(t)| \xrightarrow{p} 0$. Then*

$$\widehat{\Lambda}_j(t) \xrightarrow{p} \Lambda_j(t), \quad \widehat{S}_j(t) \xrightarrow{p} S_j(t).$$

Proofs of this theorem (and similar versions for survival data) can be found in the literature [11, 42, 43], and are not repeated here. We do note that the proofs are invariant to the method of estimation of $K_i(t)$ and when Aalen's linear model is used, invariant to the type of generalized inverse of the matrix $\mathbf{A}(t)$ selected. These choices do play a role in defining the asymptotic distribution of $\widehat{\Lambda}_j(t)$, as will be seen below.

The proof of the consistency of $\widehat{\Lambda}_j(t)$ and $\widehat{S}_j(t)$ rests on a martingale representation for $\widehat{\Lambda}_j(t)$ developed by Satten and Datta [44], which we establish in the following lemma.

Lemma II.4 *Let $J_j^*(t) = I[Y_j^*(t) > 0]$. For $\widehat{\Lambda}_j(t)$ defined in (2.2),*

$$\begin{aligned} \widehat{\Lambda}_j(t) - \Lambda_j(t) &= \int_0^t \frac{J_j^*(s)}{Y_j^*(s)} dM_j^*(s) + \\ &\quad \int_0^\infty \zeta_j^T(s, t) [\mathbf{I} - \widetilde{\mathbf{X}}(s)\mathbf{A}^{-1}(s)\widetilde{\mathbf{X}}^T(s)] d\mathbf{M}^c(s) + o_p(n^{-1/2}), \end{aligned} \tag{2.3}$$

where

$$\widetilde{\mathbf{X}}(s) = \begin{pmatrix} I[T_1 > s] \cdot \mathbf{X}_1^T(s) \\ \vdots \\ I[T_n > s] \cdot \mathbf{X}_n^T(s) \end{pmatrix}, \quad \mathbf{M}^c(s) = \begin{pmatrix} M_1^c(s) \\ \vdots \\ M_n^c(s) \end{pmatrix}, \quad \zeta_j(s, t) = \begin{pmatrix} \zeta_{1j}(s, t) \\ \vdots \\ \zeta_{nj}(s, t) \end{pmatrix},$$

$$\begin{aligned} \zeta_{ij}(s, t) &= \frac{\Phi_j}{nK_i(s)} \left(\frac{I[s - t < T_{ij}^*]}{S_j(t \wedge (U_{ij}^* - T_{ij}^*))} - \frac{I[s - t < T_{ij}^*]}{S_j((0 \vee (s - T_{ij}^*)) \wedge (U_{ij}^* - T_{ij}^*))} \right. \\ &\quad \left. - \frac{I[s < U_{ij}^*, U_{ij}^* - T_{ij}^* \leq t]}{S_j(U_{ij}^* - T_{ij}^*)} \right), \end{aligned} \tag{2.4}$$

and $\Phi_j = P[X_{ij} = 1]$.

The proof of Lemma II.4 is long and fairly technical, and we defer it to the Appendix. We note that this expression corrects errors in the limits of integration in the second integral of (2.3) and in the definition of $\zeta_{ij}(s, t)$ in (2.4) defined by Satten and Datta [44].

The martingale representation (2.3) provides an avenue for deriving the asymptotic distribution of $\widehat{\Lambda}_j(t)$ and $\widehat{S}_j(t)$ via the martingale central limit theorem (cf. Andersen, et al [6], Thm. II.5.1). As our focus is on K -sample test statistics and nonparametric regression coefficient estimators, we do not further explore the asymptotic properties of $\widehat{\Lambda}_j(t)$ and $\widehat{S}_j(t)$. We do, however, remark that (II.4) is the sum of two martingales which are orthogonal with respect to the filtration generated by the observed data, $\sigma(\mathcal{F}_t, \overline{\mathbf{X}}_i, 1 \leq i \leq n)$:

1. a martingale derived as the stochastic integral of a predictable process with respect to the waiting time martingale – $M_j^*(t)$,
2. a martingale derived as the stochastic integral of a predictable process with respect to the vector-valued censoring hazard martingale – $\mathbf{M}^c(t)$.

The orthogonality of these martingales is important – the asymptotic properties of $\widehat{\Lambda}_j$ consequently can be derived from those of the two component martingales with no consideration of their covariance. This martingale representation will be revisited in defining the asymptotic properties of our K -sample test statistics and nonparametric regression estimators for waiting times in the following chapters.

CHAPTER III

NONPARAMETRIC K -GROUP HYPOTHESIS TESTING

We now turn our attention to the comparison of stage j waiting time distributions among K groups of individuals. The null hypothesis is simply stated, $H_0 : \Lambda_{1j} = \cdots = \Lambda_{Kj}$, where Λ_{hj} denotes the stage j cumulative hazard function for group h . We note that H_0 is equivalent to hypothesizing the equivalence of the K waiting time distributions, given that we have assumed continuous waiting time distributions and the functional equivalence $\Lambda_{hj}(t) = -\log S_{hj}(t)$. Many well-known test statistics for survival data, including the log-rank test and the more general class of Fleming-Harrington [16] tests, can be derived as the weighted accumulation of differences between estimates of the K group-specific cumulative hazards and the overall cumulative hazard [6]. We develop our waiting time test statistics analogously. As with the Satten and Datta [44] cumulative hazard and survival function estimators for waiting times described in the previous chapter, we proceed by introducing test statistics for uncensored waiting time data (adapted from test statistics for survival data) and then replacing the uncensored data counting process composing these statistics with their IPCW counterparts.

We begin the chapter with a brief exposition of the extension of our notation and the affirmation of the results from the previous chapter to the K -sample setting. We then introduce test statistics for uncensored waiting time data as described by Andersen, et al [6] for survival data. We then derive our IPCW test statistics and establish their asymptotic normality and then proceed to an exploration of their empirical size and power via simulation studies for waiting time data. In the validation portion of our simulation study, we assess the properties of our test while noting the inadequacy of traditional survival data test statistics under violations of the semi-Markov property. In the comparison portion of our simulation study, we compare the power of our test with a Mann-Whitney type test for waiting times [15] for proportional hazards alternatives, for which the log rank test for survival data is most powerful. We conclude the chapter by illustrating the use of our test statistics with an analysis of a data set of 154 burn patients [23].

Expansion of Notation

In general, we will prefix the subscript h ($1 \leq h \leq K$) to subscripts i , j , and ij to denote group membership. The index i for individuals consequently ranges from 1 to n_h , where n_h denotes the number of individuals in group h , and $\sum_h n_h = n$. The stage index j is unaffected. We summarize the notational expansion in the following list.

- The uncensored data quantities T_{hij}^* , U_{hij}^* , T_{hi}^* , and X_{hi} and the censoring times C_{hi} are now indexed for subject i in group h , and are otherwise defined identically as before.
- The censored data quantities T_{hij} , U_{hij} , T_{hi} , γ_{hij} , and δ_{hij} and the covariate vectors \mathbf{X}_{hi} and histories $\bar{\mathbf{X}}_{hi}$ are now indexed for subject i in group h , and are otherwise defined identically as before. The history of transition times \mathcal{F}_t is pooled over the K groups and remains unchanged.
- All counting processes – $N_{hj}^*(t)$, $N_{hj}(t)$, \bar{N}_{hj} , \widehat{N}_{hj} , $Y_{hj}^*(t)$, $Y_{hj}(t)$, \bar{Y}_{hj} , and \widehat{Y}_{hj} – are now indexed for group h and are otherwise defined identically as before. These counting processes pooled over the K groups will be denoted as $N_{\cdot j}^*(t) = \sum_h N_{hj}^*(t)$, $N_{\cdot j}(t) = \sum_h N_{hj}(t)$, etc.
- The cumulative hazard and survival functions $\Lambda_{hj}(t)$ and $S_{hj}(t)$ are now indexed for group h , as are their uncensored data estimators $\Lambda_{hj}^*(t)$ and $S_{hj}^*(t)$ and IPCW estimators $\widehat{\Lambda}_{hj}(t)$ and $\widehat{S}_{hj}(t)$. The estimators are defined in terms of the group specific counting processes, but are otherwise defined identically as before. The corresponding quantities pooled over the K groups are denoted $\Lambda_{\cdot j}(t)$, $S_{\cdot j}(t)$, $\Lambda_{\cdot j}^*(t)$, $S_{\cdot j}^*(t)$, $\widehat{\Lambda}_{\cdot j}(t)$, and $\widehat{S}_{\cdot j}(t)$, with the estimators defined in terms of the pooled counting processes, (e.g.) $\Lambda_{hj}^*(t) = \int_0^t dN_{hj}^*(s)/Y_{hj}^*(s)$.
- The censoring hazard $\lambda_{hi}^c(t)$ and corresponding product integral $K_{hi}(t)$ are now indexed for subject i in group h , and are otherwise defined identically as before, as are the counting processes for censoring events $N_{hi}^c(t)$ and the associated martingale $M_{hi}^c(t)$.
- Aalen's linear model is defined as $\lambda_{hi}^c(t) = \beta_0^c(t) + \sum_{m=1}^p \beta_m^c(t) X_{him}(t)$ for individual i in group h . The integrated coefficient vector $\mathbf{B}^c(t)$ is estimated by $\widehat{\mathbf{B}}^c(t) = \sum_{h=1}^K \sum_{i=1}^{n_h} I(T_{hi} \leq t) (1 - \delta_{hi}) \mathbf{A}^{-1}(T_{hi}) \mathbf{X}_{hi}(T_{hi})$. The cumulative hazard of

censoring and product integral estimators are $\widehat{\Lambda}_{hi}^c(t) = \int_0^t \mathbf{X}_{hi}^T(s) d\widehat{\mathbf{B}}^c(s)$ and $\widehat{K}_{hi}^c(t) = \prod_{s \leq t} (1 - d\widehat{\Lambda}_{hi}^c(s))$, analogous to the 1-sample setting with the added subscripts h . Note that $\widehat{\mathbf{B}}^c(t)$ is pooled over the K groups; group-specific estimates can be obtained by explicitly including a $K - 1$ vector of covariates indicating group status.

- Group-specific versions of Theorem II.1, Lemma II.2, Theorem II.3, and Lemma II.4 hold with proofs identical to the one-sample setting. In Lemma II.4, the matrix $\widetilde{\mathbf{X}}(s)$ and the vector $\mathbf{M}^c(s)$ are ordered lexicographically, as is the modified vector $\zeta_j^K(s, t)$, as follows:

$$\widetilde{\mathbf{X}}(s) = \begin{pmatrix} I[\mathbf{T}_{11} > s] \widetilde{\mathbf{X}}_{11}^T(s) \\ \vdots \\ I[\mathbf{T}_{1n_1} > s] \widetilde{\mathbf{X}}_{1n_1}^T(s) \\ \vdots \\ I[\mathbf{T}_{K1} > s] \widetilde{\mathbf{X}}_{K1}^T(s) \\ \vdots \\ I[\mathbf{T}_{Kn_K} > s] \widetilde{\mathbf{X}}_{Kn_K}^T(s) \end{pmatrix}, \quad \mathbf{M}^c(s) = \begin{pmatrix} M_{11}^c(s) \\ \vdots \\ M_{1n_1}^c(s) \\ \vdots \\ M_{K1}^c(s) \\ \vdots \\ M_{Kn_K}^c(s) \end{pmatrix}, \quad \zeta_j^K(s, t) = \begin{pmatrix} \zeta_{11j}^K(s, t) \\ \vdots \\ \zeta_{1n_1j}^K(s, t) \\ \vdots \\ \zeta_{K1j}^K(s, t) \\ \vdots \\ \zeta_{Kn_Kj}^K(s, t) \end{pmatrix},$$

where the functions $\zeta_{hij}^K(s, t)$ are modifications of the functions $\zeta_{ij}(s, t)$ as defined in (2.4):

$$\zeta_{hij}^K(s, t) = \frac{\Phi_{hj}}{n_h K_{hi}(s)} \left(\frac{I[s - t < T_{hij}^*]}{S_{hj}(t \wedge (U_{hij}^* - T_{hij}^*))} - \frac{I[s - t < T_{hij}^*]}{S_{hj}((0 \vee (s - T_{hij}^*)) \wedge (U_{hij}^* - T_{hij}^*))} - \frac{I[s < U_{hij}^*, U_{hij}^* - T_{hij}^* \leq t]}{S_{hj}(U_{hij}^* - T_{hij}^*)} \right),$$

where $\Phi_{hj} = P[X_{hij} = 1]$ denotes the probability of an individual in group h ever reaching stage j . Note that the superscript K in $\zeta_j^K(s, t)$ is not an index, rather meant to denote the K -sample version the function $\zeta_j(s, t)$ defined in (2.4). Further, we extend the martingale representation of $\widehat{\Lambda}_j(t)$ in (2.3) to the K -sample setting as follows:

$$\begin{aligned} \widehat{\Lambda}_{hj}(t) - \Lambda_{hj}(t) &= \int_0^t \frac{J_{hj}^*(s)}{Y_{hj}^*(s)} dM_{hj}^*(s) + \\ &\int_0^\infty \zeta_j^K(s, t)^T [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s)] d\mathbf{M}^c(s) + o_p(n^{-1/2}), \end{aligned} \tag{3.1}$$

We preserve our notational conventions with respect to the quantities and functions of interest. Uncensored data quantities (entry and exit times) are marked with asterisks and censored data quantities are not marked. Censored data functions (counting processes, covariate functions) are marked with asterisks, censored data functions are unmarked, and weighted functions (counting processes, estimators) are hatted. With this K -sample notation and group-specific results from Chapter II, we now proceed to developing our waiting time test statistics.

Waiting Time Test Statistics for K Groups

In order to develop our test statistics as weighted accumulations of differences between the group-specific and pooled cumulative hazard estimates, we define the stochastic weight processes $W_{hj}^*(t)$, $1 \leq h \leq K$. As in Andersen, et al [6], we constrain attention to weight processes of the form $W_{hj}^*(t) = W_j^*(t)Y_{hj}^*(t)$, where $W_j^*(t)$ is a non-negative, locally bounded, predictable function depending only on the pair $(N_j^*(t), Y_j^*(t))$. This general formulation covers many test statistics for survival data, including the log-rank and Fleming-Harrington class of tests, although more general weight processes are permitted.

To test $H_0 : \Lambda_{1j} = \dots = \Lambda_{Kj}$ for uncensored waiting times, we define the test statistics

$$Z_{hj}^*(t) = \int_0^t W_j^*(s)Y_{hj}^*(s)d(\Lambda_{hj}^* - \Lambda_j^*)(s), \quad 1 \leq h \leq K. \quad (3.2)$$

The integral in (3.2) accumulates the differences between the estimated group-specific cumulative hazard, $\Lambda_{hj}^*(t) = \int_0^t dN_{hj}^*(s)/Y_{hj}^*(s)$, and the pooled cumulative hazard, $\Lambda_j^*(t) = \int_0^t dN_j^*(s)/Y_j^*(s)$, weighted by $W_j^*(s)Y_{hj}^*(s)$. The statistics $Z_{hj}^*(t)$ present a reasonable test of H_0 at face – large deviations of group-specific cumulative hazards from the pooled cumulative hazards produce large values of $Z_{hj}^*(t)$ and provide evidence against H_0 . The weight process provides some flexibility in tailoring $Z_{hj}^*(t)$ to a given data set or a priori suspicion about the nature of the difference in hazard between groups. For example, weight processes can be defined that more heavily weight early differences in the cumulative hazard functions.

Plugging in the Nelson-Aalen estimators Λ_{hj}^* and Λ_j^* into (3.2) produces

$$Z_{hj}^*(t) = \int_0^t W_j^*(s)dN_{hj}^*(s) - \int_0^t W_j^*(s)\frac{Y_{hj}^*(s)}{Y_j^*(s)}dN_j(s).$$

Under the null hypothesis, $\Lambda_{hj}^* = \Lambda_j^*$, so we can add zero in the form of $d\Lambda_j^*(s) - d\Lambda_{hj}^*(s)$ to get

$$\begin{aligned} Z_{hj}^*(t) &= \int_0^t W_j^*(s) d(N_{hj}^* - \Lambda_{hj}^*)(s) - \int_0^t W_j^*(s) \frac{Y_{hj}^*(s)}{Y_j^*(s)} d(N_j - \Lambda_j^*)(s) \\ &= \int_0^t W_j^*(s) dM_{hj}^*(s) - \int_0^t W_j^*(s) \frac{Y_{hj}^*(s)}{Y_j^*(s)} dM_j^*(s) \\ &= \sum_{l=1}^k \int_0^t W_j^*(s) \left(\Delta_{hl} - \frac{Y_{hj}^*(s)}{Y_j^*(s)} dM_{lj}^*(s) \right), \end{aligned}$$

where $M_j^* = \sum_h M_{hj}^*$ and Δ_{hl} denotes the Kronecker delta. Based on this martingale representation, the predictable covariation process of the $Z_{hj}^*(t)$ is

$$\langle Z_{hj}^*(t), Z_{kj}^*(t) \rangle = \int_0^t W_j^*(s)^2 Y_{hj}^*(s) \left(\Delta_{hk} - \frac{Y_{kj}^*(s)}{Y_j^*(s)} \right) \lambda_j(s) ds. \quad (3.3)$$

The expectation of (3.3) provides the covariance of the $Z_{hj}^*(t)$ and $Z_{kj}^*(t)$, which can be estimated by

$$[Z_{hj}^*(t), Z_{kj}^*(t)] = \int_0^t W_j^*(s)^2 \frac{Y_{hj}^*(s)}{Y_j^*(s)} \left(\Delta_{hk} - \frac{Y_{kj}^*(s)}{Y_j^*(s)} \right) dN_j^*(s). \quad (3.4)$$

Tests of H_0 can be based upon the quadratic form $\mathbf{Z}_j^*(t)^T \Sigma_j^{*-}(t) \mathbf{Z}_j^*(t)$, where the vector $\mathbf{Z}_j^*(t) = (Z_{1j}^*(t), \dots, Z_{Kj}^*(t))$, $\Sigma_j^*(t)$ is the matrix with $(h, k)^{th}$ element defined by (3.4), and $\Sigma_j^{*-}(t)$ is a generalized inverse of $\Sigma_j^*(t)$. Under mild conditions (see Andersen, et al [6]), $\Sigma_j^*(t)$ has rank $K - 1$ and an alternative but equivalent quadratic form can be constructed as $\mathbf{Z}_{0j}^*(t)^T \Sigma_{0j}^{*-1}(t) \mathbf{Z}_{0j}^*(t)$, where $\mathbf{Z}_{0j}^*(t)$ is $\mathbf{Z}_j^*(t)$ with a single element removed and $\Sigma_{0j}^*(t)$ is $\Sigma_j^*(t)$ with the corresponding row and column removed. Under mild regularity conditions, $\mathbf{Z}_j^*(t)$ converges weakly to a normal limit (cf. Andersen, et al [6], Thm. V.2.1), so that either of the above quadratic forms follows the χ^2 distribution, providing critical values for the test statistic.

Clearly, $\mathbf{Z}_j^*(t)$ is of little use for right-censored multi-stage data being based on uncensored data quantities. Further, test statistics based on unweighted, censored data counting processes would require independent censoring and the semi-Markov property to be valid. Hence, we propose testing H_0 with IPCW versions of (3.2) where, as with the Satten and Datta [44] estimators described in Chapter II, we replace uncensored data quantities with their IPCW counterparts. We can then define IPCW K -sample test

statistics for waiting times in a multi-stage model as,

$$\widehat{Z}_{hj}(t) = \int_0^t \widehat{W}_j(s) \widehat{Y}_{hj}(s) d(\widehat{\Lambda}_{hj} - \widehat{\Lambda}_{\cdot j})(s), \quad 1 \leq h \leq K, \quad (3.5)$$

where $\widehat{\Lambda}_{hj}(t)$ and $\widehat{\Lambda}_{\cdot j}(t)$ are as defined in (2.2) in Chapter II. Note that we have redefined the weight process $\widehat{W}_j(s) \widehat{Y}_{hj}(s)$ in terms of the IPCW counting process $\widehat{Y}_{hj}(s)$, where $\widehat{W}_j(s)$ remains a non-negative, locally bounded, predictable function but now depends on the IPCW pair $(\widehat{N}_{\cdot j}(t), \widehat{Y}_{\cdot j}(t))$. We will generally focus on the log-rank weight process $\widehat{W}_j(s) = \widehat{J}_j(s) = I[\widehat{Y}_j(s) > 0]$. In particular, we will use this weight process in our simulation study of $\widehat{Z}_{hj}(t)$. The formulation of $\widehat{Z}_{hj}(t)$ in (3.5) can alternatively be expressed as

$$\widehat{Z}_{hj}(t) = \int_0^t \widehat{W}_j(s) d\widehat{N}_{hj}(s) - \int_0^t \widehat{W}_j(s) \frac{\widehat{Y}_{hj}(s)}{\widehat{Y}_{\cdot j}(s)} d\widehat{N}_{\cdot j}(s).$$

We note that this formulation highlights that $\widehat{Z}_{hj}(t)$ also accumulates differences between what is observed for group h and what is expected. Temporarily ignoring the weights $\widehat{W}_j(t)$, the first term merely counts the group h exits from stage j , while the second term counts exits from stage j for all groups weighted by the IPCW-estimated proportion of subjects in group $h - \widehat{Y}_{hj}(s)/\widehat{Y}_{\cdot j}(s)$.

Asymptotic Properties of the Test Statistic

The use of $Z_{hj}(t)$ in testing requires derivation of the asymptotic distribution of the vector $\widehat{\mathbf{Z}}_j(t) = (Z_{1j}(t), \dots, Z_{Kj}(t))$. We begin by deriving a martingale representation for the $Z_{hj}(t)$ which follows from the representation for $\widehat{\Lambda}_{hj}(t)$ in (3.1). Under the null hypothesis, note that $\Lambda_{hj} = \Lambda_{\cdot j}$, and hence for $1 \leq h \leq K$,

$$\widehat{Z}_{hj}(t) = \int_0^t \widehat{W}_j(s) \widehat{Y}_{hj}(s) d(\widehat{\Lambda}_{hj} - \Lambda_{hj})(s) - \int_0^t \widehat{W}_j(s) \widehat{Y}_{hj}(s) d(\widehat{\Lambda}_{\cdot j} - \Lambda_{\cdot j})(s).$$

Substituting in the martingale representations for $\widehat{\Lambda}_{hj}(t) - \Lambda_{hj}(t)$ and $\widehat{\Lambda}_{\cdot j}(t) - \Lambda_{\cdot j}(t)$ gives

$$\begin{aligned} \widehat{Z}_{hj}(t) &= \int_0^t \widehat{W}_j(s) \widehat{Y}_{hj}(s) \frac{J_{hj}^*(s)}{Y_{hj}^*(s)} dM_{hj}^*(s) - \int_0^t \widehat{W}_j(s) \widehat{Y}_{hj}(s) \frac{J_{\cdot j}^*(s)}{Y_{\cdot j}^*(s)} dM_{\cdot j}^*(s) - \\ &\int_0^\infty \left\{ \int_0^t \widehat{W}_j(u) \widehat{Y}_{hj}(u) \zeta_j^K(s, du)^\top \right\} [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^\top(s)] d\mathbf{M}^c(s) + \\ &\int_0^\infty \left\{ \int_0^t \widehat{W}_j(u) \widehat{Y}_{hj}(u) \zeta_j(s, du)^\top \right\} [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^\top(s)] d\mathbf{M}^c(s) + o_p(n^{-1/2}). \end{aligned} \quad (3.6)$$

Note that the first and third terms of (3.6) are the martingale expression for $\widehat{\Lambda}_{hj}(t)$ and the second and fourth terms for $\widehat{\Lambda}_j(t)$. By exploiting the asymptotic equivalence of $Y_j^*(t)$ and $\widehat{Y}_j(t)$ (noted in the proof of Lemma II.4) and noting that $M_j^* = \sum_h M_{hj}^*$, we can further simplify (3.6):

$$\begin{aligned} \widehat{Z}_{hj}(t) &= \sum_{k=1}^K \int_0^t \widehat{W}_j(s) \left(\Delta_{hk} - \frac{Y_{kj}^*(s)}{Y_j^*(s)} \right) dM_{kj}^*(s) + \\ &\quad \int_0^\infty \boldsymbol{\eta}_j^K(s, t)^\top \left[\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^\top(s) \right] d\mathbf{M}^c(s) + o_p(n^{-1/2}), \end{aligned} \quad (3.7)$$

where $\boldsymbol{\eta}_j^K(s, t)$ is defined by

$$\boldsymbol{\eta}_j^K(s, t) = \int_0^t \widehat{W}_j(u) \widehat{Y}_{hj}(u) \left(\boldsymbol{\zeta}_j^K(s, du) - \boldsymbol{\zeta}_j(s, du) \right).$$

As with the expansion of $\widehat{\Lambda}_{hj}(t)$, the martingale representation for $\widehat{Z}_{hj}(t)$ is the sum of two martingales, one associated with the stage j exit counting process ($M_{hj}^*(t)$) and one with the counting process of censoring events ($\mathbf{M}^c(t)$). Further, the martingales are orthogonal with respect to the filtration of the observed data. Hence, the asymptotic properties of the K -vector of test statistics $\widehat{\mathbf{Z}}_j(t)$ can be derived from the asymptotic properties of the two component martingales via the martingale central limit theorem, and covariance considerations can be ignored. It is also noteworthy that the first term of (3.7) directly corresponds to the uncensored data K -group test statistic in (3.2), for which conditions required for asymptotic normality are well-defined [6].

An immediate consequence of the orthogonality of the two martingales in (3.7) is that the predictable covariation process (cf. Eq. 2.3.7, [6]) of two components of $\widehat{\mathbf{Z}}_j(t)$, say $\widehat{Z}_{hj}(t)$ and $\widehat{Z}_{kj}(t)$, can be defined as the sum of the predictable covariation processes of the component terms of (3.7). Specifically,

$$\begin{aligned} \langle \widehat{Z}_{hj}(t), \widehat{Z}_{kj}(t) \rangle &= \int_0^t \widehat{W}_j^2(s) Y_{hj}^*(s) \left(\Delta_{hk} - \frac{Y_{kj}^*(s)}{Y_j^*(s)} \right) \lambda_j(s) ds + \\ &\quad \int_0^\infty \boldsymbol{\eta}_j^K(s, t)^\top \left[\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^\top(s) \right] \times \mathbf{I}_\lambda(s) \times \\ &\quad \left[\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^\top(s) \right] \boldsymbol{\eta}_j^K(s, t) ds + o_p(n^{-1/2}), \end{aligned} \quad (3.8)$$

where $\mathbf{I}_\lambda(s)$ is the matrix $\mathbf{diag}\{I[T_{11} \geq s] \lambda_{11}^c(s), \dots, I[T_{K n_K} \geq s] \lambda_{K n_K}^c(s)\}$. The asymptotic covariance of $Z_{hj}(t)$ and $Z_{kj}(t)$ is then the in-probability limit of (3.8). Let $\sigma_{hk}^2(t)$ denote

the in-probability limit of the predictable covariation process (3.8) and define the matrix $\Sigma_j(t) = \{\sigma_{hk}^2(t)\}$. Other than establishing a limiting variance for $\widehat{\mathbf{Z}}_j(t)$, this expression is of little practical use as it based upon the uncensored data counting processes $Y_{hj}^*(t)$ and $Y_{.j}^*(t)$, the population quantities $\lambda_{.j}(s)$ and $\lambda_{hi}^c(s)$, and $\zeta_j^K(s, t)$, which is constructed from the uncensored entry and exit times T_{hij}^* and U_{hij}^* and the population quantities $S_{hj}(t)$ and Φ_{hj} . We can estimate (3.8) by the optional covariation process (cf. Eq. 2.3.8, [6]) of $\widehat{\mathbf{Z}}_{hj}(t)$ and $\widehat{\mathbf{Z}}_{kj}(t)$ as follows:

$$\begin{aligned} [\widehat{\mathbf{Z}}_{hj}(t), \widehat{\mathbf{Z}}_{kj}(t)] &= \int_0^t \widehat{W}_j^2(s) \frac{\widehat{Y}_{hj}(s)}{\widehat{Y}_{.j}(s)} \left(\Delta_{hk} - \frac{\widehat{Y}_{kj}(s)}{\widehat{Y}_{.j}(s)} \right) d\widehat{N}_{.j}(s) + \\ &\quad \int_0^\infty \widehat{\eta}_j^K(s, t)^T \left[\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s) \right] \times \mathbf{I}_N(s) \times \\ &\quad \left[\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s) \right] \widehat{\eta}_j^K(s, t), \end{aligned} \quad (3.9)$$

where $\widehat{\eta}_j^h(s, t) = \int_0^t \widehat{W}_j(u) \widehat{Y}_{hj}(u) \left(\widehat{\zeta}_j^K(s, du) - \widehat{\zeta}_j(s, du) \right)$, $\mathbf{I}_N(s)$ denotes the diagonal matrix $\text{diag}\{dN_{11}^c(s), \dots, dN_{K n_K}^c(s)\}$, and the components of $\widehat{\zeta}_j^K(s, t)$ are

$$\begin{aligned} \widehat{\zeta}_{hij}^K(s, t) &= \frac{\widehat{\Phi}_{hj}}{n_h \widehat{K}_{hi}(s)} \Delta_{hk} \left(\frac{I[s - t < T_{hij}]}{\widehat{S}_{hj}(t \wedge (U_{hij} - T_{hij}))} - \right. \\ &\quad \left. \frac{I[s - t < T_{hij}]}{\widehat{S}_{hj}((0 \vee (s - T_{hij})) \wedge (U_{hij} - T_{hij}))} - \frac{I[s < U_{hij}; U_{hij} - T_{hij} \leq t]}{\widehat{S}_{hj}(U_{hij} - T_{hij})} \right). \end{aligned} \quad (3.10)$$

Note also that $\widehat{\zeta}_{hij}^K(s, t) = \frac{n_h \Phi_{.j}}{n \Phi_{hj}} \widehat{\zeta}_{hij}(s, t)$ to complete the definition of the optional variation process (3.9). In (3.10), $\widehat{S}_{hj}(t)$ represents the Satten-Datta estimator and the estimator $\widehat{\Phi}_{hj}$ can be calculated using the method of Datta and Satten [11]. Denote the optional variation process in (3.9) by $\widehat{\sigma}_{hk}^2(t)$ and define the matrix $\widehat{\Sigma}_j(t) = \{\widehat{\sigma}_{hk}^2(t)\}$.

Given the variance expressions (3.8) and (3.9) and that $\widehat{\mathbf{Z}}_j(t)$ is a martingale, being the sum of two orthogonal martingales, we can assert the asymptotic normality of $\widehat{\mathbf{Z}}_j(t)$ by imposing the conditions of Rebelledo's martingale central limit theorem. We note that the two assumptions required for a martingale to be asymptotically normal (c.f. Thm. II.5.1, [6]) are that (1) either its predictable or optional covariation process converges in probability to a continuous, deterministic, positive semi-definite limit with positive semi-definite increments in time, and (2) the Lindeberg condition, that its "path" in time approaches a continuous limit. The following theorem establishes our main result for the test statistic $\mathbf{Z}_j(t)$.

Theorem III.1 Under $H_0 : \Lambda_{1j} = \dots = \Lambda_{Kj}$ and suitable regularity conditions,

$$\sqrt{n}\widehat{\mathbf{Z}}_j(t) \xrightarrow{\mathcal{D}} \mathbf{U}(t),$$

as $n \rightarrow \infty$ where $\mathbf{U}(t) = (U_1(t), \dots, U_K(t))$ is a Gaussian martingale with covariance matrix $\Sigma_j(t) = \{\sigma_{hk}^2(t)\}$.

The definition of the regularity conditions stated in theorem and the proof of the theorem are deferred to the Appendix. As a consequence of this result and similarly to the uncensored data test statistic, we can establish a K -group test of H_0 by calculating the χ^2 statistic $\widehat{\mathbf{Z}}_j(t)\widehat{\Sigma}_j^-(t)\widehat{\mathbf{Z}}_j(t)$, where $\widehat{\Sigma}_j^-(t)$ is a generalized inverse of $\widehat{\Sigma}(t)$. Again, under mild conditions (see, e.g., [6]), this statistic will have rank $K - 1$, and tests of H_0 can alternatively be based upon the chi-square statistic $\widehat{\mathbf{Z}}_{0j}(t)\widehat{\Sigma}_0^{-1}(t)\widehat{\mathbf{Z}}_{0j}(t)$, where $\widehat{\mathbf{Z}}_{0j}(t)$ is $\widehat{\mathbf{Z}}_j(t)$ with a single row removed, $\widehat{\Sigma}_0(t)$ is $\widehat{\Sigma}(t)$ with the corresponding row and column removed, and $\widehat{\Sigma}_0^{-1}(t)$ is an ordinary inverse of $\widehat{\Sigma}_0(t)$. We demonstrate the use of this chi-square statistic in the following sections.

Simulation Study – Validation

In this section, we present the results of a simple simulation study that examined the size and power of our proposed test statistic. We compared the size and power of the log-rank version of $\widehat{\mathbf{Z}}_j(t)$, for which $\widehat{W}_j(t) = \widehat{J}_j(t) = I[\widehat{Y}_j(t) > 0]$, to that of the log-rank test statistic for survival data adapted to stage j waiting times, defined as the vector $\mathbf{Z}_j(t) = (Z_{1j}(t), \dots, Z_{Kj}(t))^T$. The components of $\mathbf{Z}_j(t)$ are defined in terms of the censored, unweighted counting processes $Y_{hj}(t)$ and $N_{hj}(t)$ defined in Chapter II as $Z_{hj}(t) = N_{hj}(t) - \int_0^t Y_{hj}(s)/Y_{.j}(s)dN_{.j}(s)$. We considered a three-stage tracking model, in which individuals entered root stage 0 at time 0 and then progressed through transient stage 1 on to absorbing stage 2 or were censored, and analyzed stage 1 waiting times. The power of $\widehat{\mathbf{Z}}_1(t)$ and $\mathbf{Z}_1(t)$ were calculated as the observed proportion of 10,000 Monte Carlo replications in which the null hypothesis of the equality of stage 1 waiting time distributions was rejected. Rejection was defined as the test statistic $\widehat{\mathbf{Z}}_{0j}(t)\widehat{\Sigma}_0^{-1}(t)\widehat{\mathbf{Z}}_{0j}(t)$ exceeding 95th percentile of the χ_1^2 distribution, where $\widehat{\mathbf{Z}}_{0j}(t)$ and $\widehat{\Sigma}_0^{-1}(t)$ were as defined in the previous section. For computational convenience, bootstrap resampling [14] at 250 loops was used to calculate the standard error of our test statistic, as the covariance

estimator (3.8) is computationally intensive, requiring in particular the double integration of the bivariate function $\hat{\eta}_j^K(s, t)$. We considered per-group sample sizes of 50, 100, and 250 for two designs under which we altered parameters governing the bivariate distribution of stage 0 and 1 waiting times as follows.

1. *Correlated lognormal stage 0 and 1 waiting times.* Let (W_0, W_1) denote the bivariate pair of stage 0 and 1 waiting times. We simulated $(\log(W_0), \log(W_1))$ as a bivariate normal pair, with mean vector $(1, 1)$ in group 1 and $(1, \mu)$ in group 2, where μ varied from 1 to 2.5 to cover the null ($\mu = 1$) and alternative ($\mu > 1$) hypotheses. Marginal variances were set equal to 1, and the correlation between $\log(W_0)$ and $\log(W_1)$ was set equal to ρ , where ρ took values -0.5, 0, and 0.5. The bivariate pair was then exponentiated to produce lognormal waiting times. We crossed the three possible correlation values between the two groups, for a total of nine variants of this design.
2. *Markov chain with gamma-distributed waiting times.* We simulated Stage 0 waiting times from the gamma distribution with shape parameter 2 and scale parameter 1 or 2, different in each group so that one group experienced longer waits (scale = 2) in stage 0. We generated stage 1 waiting times that were positively and negatively associated with stage 0 waiting times through the equations

$$F_{\Gamma}^{-1}(U * F_{\Gamma}(W_0)) \text{ and } F_{\Gamma}^{-1}(U * (1 - F_{\Gamma}(W_0))), \quad (3.11)$$

where W_0 represents the waiting time for stage 0, F_{Γ} the gamma distribution function with group-appropriate shape and scale parameters, U a uniform $(0, 1)$ random variable, and F_{Γ}^{-1} the gamma quantile function. In group 1, this quantile function had shape parameter 2 and scale parameter 1. In group 2, the shape parameter for the quantile function ranged from 2 to 3.5, to cover the null (shape = 2) and alternative (shape > 2) hypotheses, and the scale parameter was 1. We crossed the three design parameters – direction of group 1 association (+/-), direction of group 2 association (+/-), stage 0 gamma scale parameter (1 or 2) – to produce 8 variants of this design. We note that the approximate rank correlations between stage 0 and stage 1 waiting times under design 2 were ± 0.67 , depending on the direction of the correlation induced by the formulas in (3.11).

Table 1

Censoring rates for test statistic simulations under the null hypothesis

Design	Waiting Time Distributions			Censoring Rates (%)		
	Stage 0	Stage 1	Correlation	Stage 0	Stage 1	Never
1	log- $N(1, 1)$	log- $N(1, 1)$	-0.5	39	39	22
			0	39	33	28
			0.5	39	28	33
2	$\Gamma(2, 1)$	$\Gamma(2, 1)$	+	18	13	69
			-	18	13	69
	$\Gamma(2, 2)$		+	42	10	48
			-	42	14	44

For each design, censoring times were generated from the Weibull distribution with shape parameter 2 and scale parameter 5, independent of stage waiting times. Consequently, to estimate $K_{hi}(t)$ and construct the weighted counting processes $\widehat{N}_{hj}(t)$ and $\widehat{Y}_{hj}(t)$, we applied Aalen's linear model with a single time-invariant covariate denoting group status, for which $\widehat{K}_{hi}(t)$ is the Kaplan-Meier estimator within each group. We selected independent censoring so that are simulations modeled a problem unique to multi-stage models, (i.e.) the data generated satisfy the independent censoring assumption but do not have the semi-Markov property. Table 1 provides the censoring rates for different combinations of the design parameters for designs 1 and 2. Note that in variants 1-4 of design 2, stage 0 waiting times from group 2 were longer (generated from the $\Gamma(2, 2)$ distribution) than waiting times from group 1 (generated from the $\Gamma(2, 1)$ distribution), and vice versa for variants 5-8.

The empirical size of our test statistic is compared to that of the naive log-rank test in Table 2. Note that given 10,000 iterations per estimated empirical size, the standard error for each calculated empirical size was no more than 0.005. For size estimates near the nominal 5% level, the standard error was approximately 0.002. Our test exhibited size close to the nominal size of 0.05, while the naive log-rank test exhibited inflated size under all variants of each design. The departure from the nominal size of 0.05 for the naive log-rank test was smallest in variants of design 1 in which the correlation between stage 0 and 1 waiting times was equal in the two groups (variants 1, 5, and 9) for $N = 100$. The empirical size of the naive log-rank test increased with sample size for all variants of both designs, and in particular for variants 1, 5, and 9 of design 1, under which the empirical

Table 2

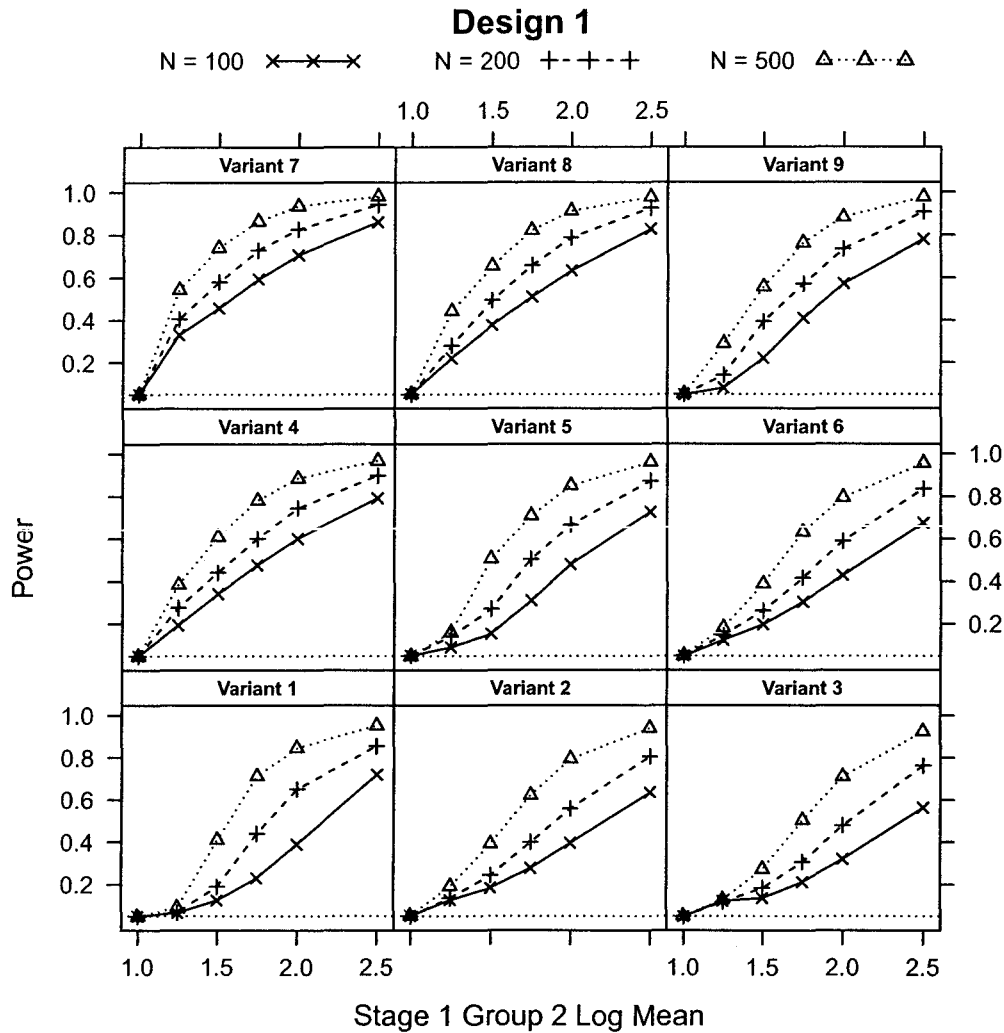
Empirical size estimates of IPCW test and naive log-rank test at nominal 5% level

Design- Variant	Waiting Time Corr.		$\hat{\mathbf{Z}}_1(t)$			$\mathbf{Z}_1(t)$		
	Group 1	Group 2	N = 100	200	500	N = 100	200	500
1-1	-0.5	-0.5	0.049	0.048	0.048	0.064	0.080	0.106
1-2		0.0	0.052	0.049	0.052	0.193	0.332	0.649
1-3		0.5	0.049	0.049	0.051	0.449	0.717	0.980
1-4	0.0	-0.5	0.049	0.048	0.048	0.093	0.142	0.284
1-5		0.0	0.051	0.050	0.048	0.057	0.056	0.059
1-6		0.5	0.050	0.051	0.052	0.145	0.224	0.490
1-7	0.5	-0.5	0.050	0.049	0.050	0.312	0.574	0.932
1-8		0.0	0.052	0.050	0.051	0.116	0.189	0.412
1-9		0.5	0.048	0.049	0.052	0.053	0.051	0.054
2-1	-	-	0.049	0.050	0.050	0.126	0.196	0.428
2-2	+	-	0.049	0.051	0.050	0.551	0.849	0.997
2-3	-	+	0.049	0.049	0.049	0.663	0.917	0.999
2-4	+	+	0.052	0.050	0.049	0.185	0.312	0.644
2-5	-	-	0.050	0.050	0.052	0.120	0.202	0.427
2-6	+	-	0.052	0.050	0.048	0.658	0.920	0.999
2-7	-	+	0.050	0.050	0.048	0.556	0.850	0.996
2-8	+	+	0.051	0.050	0.049	0.187	0.318	0.644

size of $\mathbf{Z}_1(t)$ was near 0.05 for $N = 100$. Further, as the between-group disparity in the correlation between stage 0 and 1 waiting times increased, the size of $\mathbf{Z}_1(t)$ further inflated.

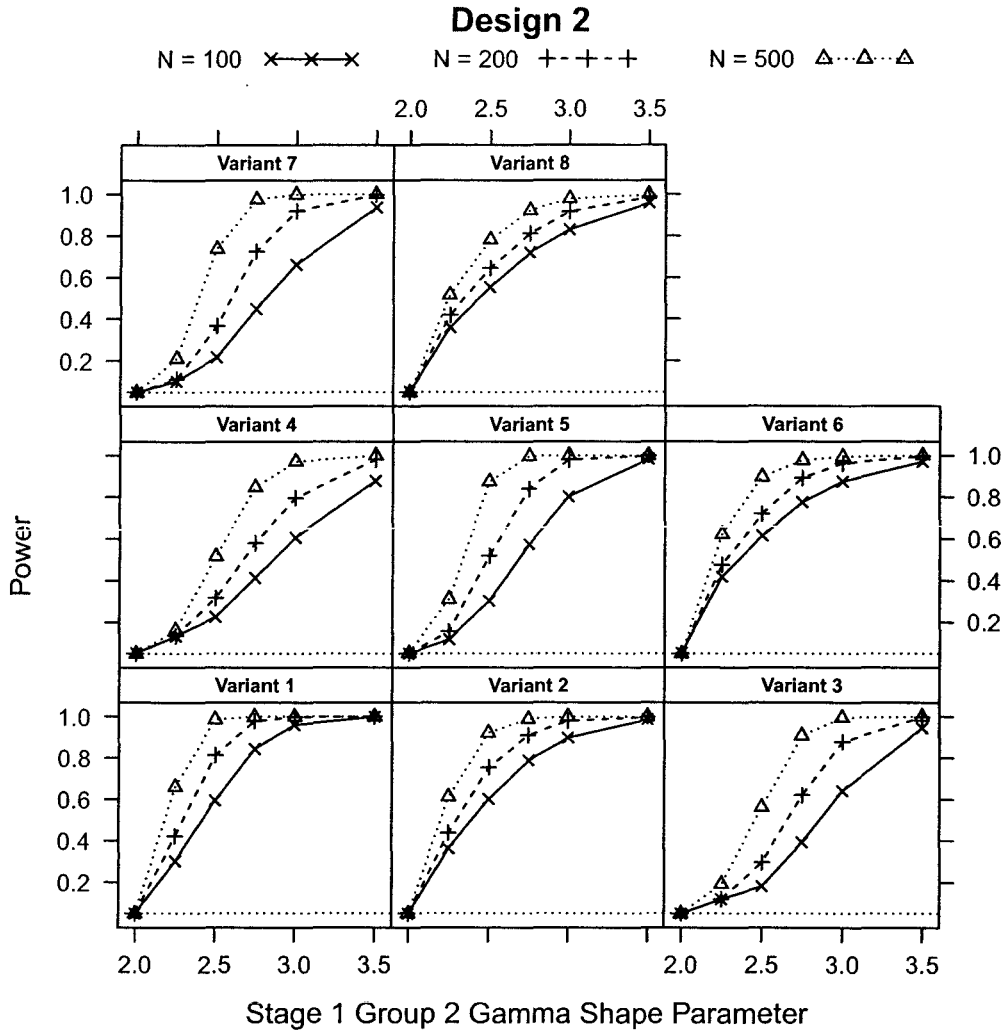
Figure 1 provides power curves for our test under design 1 of the simulation study. We noted that the power of our test statistic increased both with sample size and with the magnitude of the marginal difference between stage 1 waiting times. We consider variant 5 of design 1 as a reference case, in which waiting times were uncorrelated for both groups. Variant 3, in which waiting times were positively correlated in group 2 and negatively correlated in group 1, exhibited reduced power relative to variant 5. This was a reasonable phenomenon. Under variant 3, longer waits in stage 1 were associated with longer waits in stage 0 in group 2. Hence, those in group 2 with long waits in stage 1, (i.e.) those providing evidence against the null hypothesis, were more likely to be censored. Conversely, longer waits in stage 1 were associated with shorter waits in stage 0, so those in group 1 with long waits in stage 1, (i.e.) those providing evidence for the null hypothesis, were less likely to be censored. The converse was true for variant 7, which exhibited greater power

Figure 1. Empirical power of \hat{Z}_1 for 9 variants of simulation design 1



than the reference case. Individuals in group 2 with long stage 1 waiting had shorter stage 0 waiting times and were consequently less likely to be censored (more evidence against the null), while individuals in group 1 with long stage 1 waiting times were more likely to be censored (less evidence for the null). In general, when correlations between stage 0 and stage 1 waiting times induced censoring that provided additional evidence against the null, (i.e.) when observations in support of the alternative were more likely to be fully observed, the power of \hat{Z}_1 increased relative to the reference case. When stage 0 and stage 1 waiting time correlations left observations in support of the null were more likely to be full observed, power decreased relative to the reference case. Therefore, we concluded that while our IPCW test statistic maintained appropriate size under variants of design 1, it was

Figure 2. Empirical power of \hat{Z}_1 for 8 variants of simulation design 2



not completely immune to dependent censoring patterns, as empirical power varied as a function of the waiting time correlation.

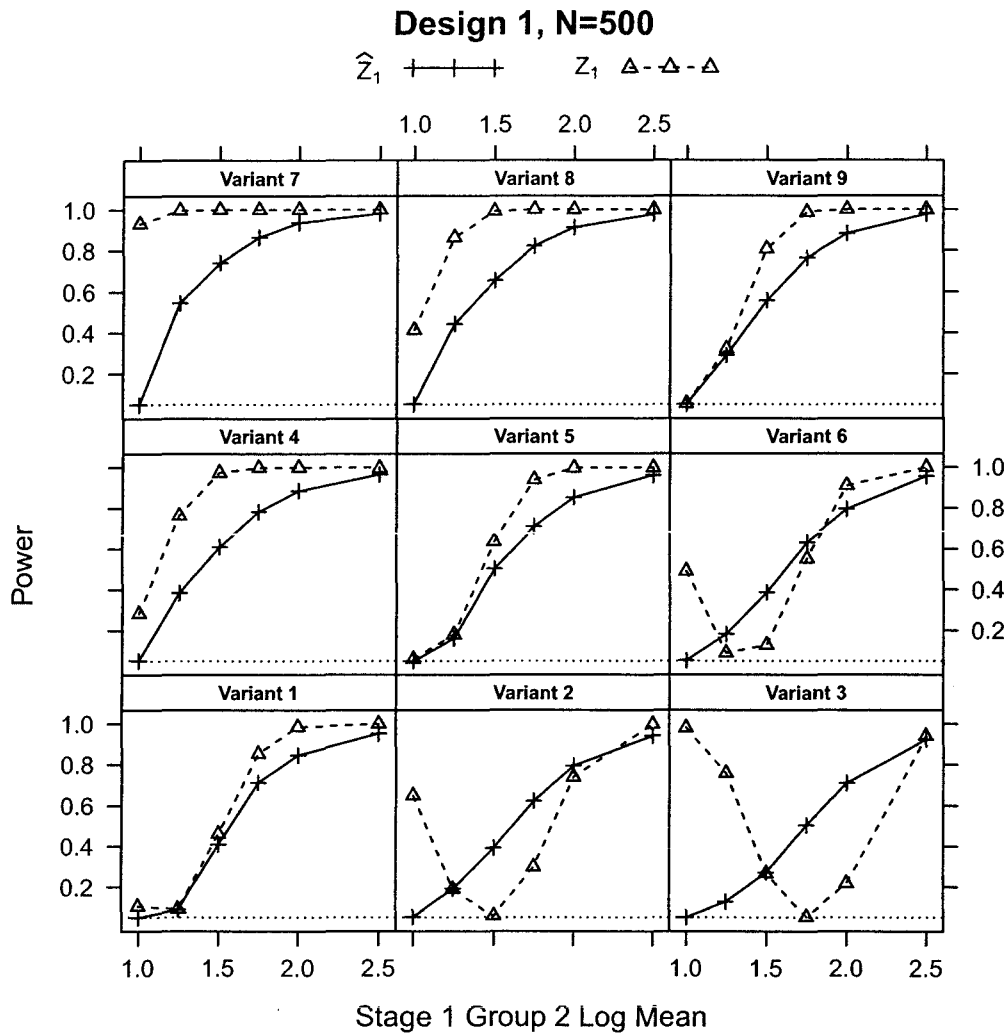
The power of \hat{Z}_1 under design 2 was more complexly related to the simulation design parameters (Figure 2), although we again noted that the power of our test increased with sample size and the marginal difference between stage 1 waiting times between groups. The dominating factor in determining the power of the test was the direction of the correlation between stage 0 and stage 1 waiting times in the group with longer stage 0 waiting times. When stage 0 waiting times were greater in group 2 (variants 1-4), negative correlation between stage 0 and stage 1 waiting times in group 2 led to higher power (variants 1 and 2). Alternatively, when stage 0 waiting times in group 1 were longer (variants 5-8), positive

correlation between stage 0 and 1 waiting times in group 1 led to higher power (variants 6 and 8). Further, the power curves for these four variants (1, 2, 6, and 8) were quite similar, indicating that other design parameters were of lesser importance. Variants 3, 4, 5, and 7, exhibited reduced power relative to the other four, and had power curves that were similar (among themselves).

The importance of the stage 0 waiting times in determining power under design 2 was not unexpected. Because censoring evolves in calendar time, longer waits prior to entry in stage 1 increase the likelihood of censoring prior to entry into or exit from stage 1 and heavier censoring tends to reduce power. Hence, any design parameter that exerted direct influence over the censoring rate, such as the stage 0 waiting times, would be expected to affect the power. The waiting time correlation, therefore, had a greater impact on power when it was altered for the group experiencing longer stage 0 waiting times. For simulations in which group 2 exhibited longer stage 0 waits, power was greatest when group 2 individuals with long stage 1 waits (supporting the alternative hypothesis) were least likely to be censored. This occurred when long stage 1 waits were associated with short stage 0 waits, (i.e.) when there were negative correlations as in variants 1 and 2. The correlation between waiting times in group 1 did not dramatically affect power since stage 0 waiting times were much less than those for group 2. Conversely, simulations in which group 1 exhibited longer stage 0 waits had greater power when group 1 individuals with long stage 1 waits (supporting the null) were most likely to be censored. This occurred when long stage 1 waits were associated with long stage 0 waits, (i.e.) when there were positive correlations as in variants 6 and 8.

Figures 3 and 4 provide comparisons of our test to the naive log-rank test under both designs with total sample size $N = 500$, and illustrate the general inappropriateness of the naive log-rank test for multi-stage data. Only for variants 1, 5, and 9 of design 1 were the power curves for the naive log-rank test sensible – power increased with the difference in marginal stage 1 waiting times, exhibited greater power than our test, and only slightly over-inflated in size. For variants 2, 3, and 6 of design 1, in which the group 2 waiting time correlation was strictly greater than the group 1 waiting time correlation, power was a non-monotonic function of the difference in marginal waiting times between groups – note that the power curves achieve minima away from the null hypothesis (1.0 on the plots). For variants 4, 7, and 8 of design 1, in which the group 2 waiting time correlation was strictly smaller than the group 1 waiting time correlation, the power curves were monotonic

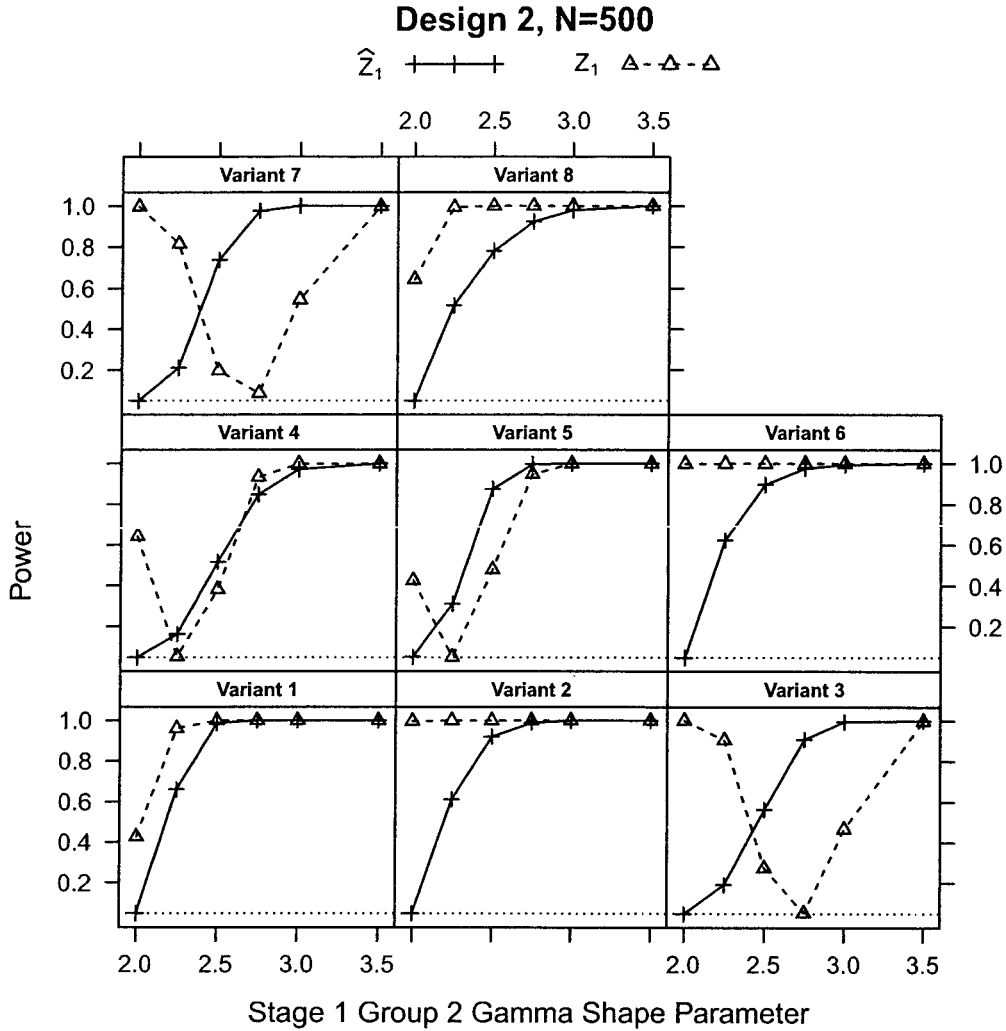
Figure 3. Power comparison of \hat{Z}_1 and Z_1 for design 1 at $N = 500$



functions of the difference in marginal waiting times between groups but exhibited greatly inflated size. For design 2, the naive log-rank test exhibited non-monotonically increasing power for variants 3, 4, 5, and 7 – the same variants under which our test exhibited reduced power. Variants 1, 2, 6, and 8, for which our test showed greater power, exhibited grossly inflated size.

From these simulations we can provisionally conclude that our test statistic \hat{Z}_1 is an appropriate K -sample test for waiting times in a multi-stage model, and that the log-rank test for survival data is not an appropriate test of waiting times from a multi-stage model. We do note that specification of the model for the censoring hazard used to generate the inverse probability of censoring weights may have an impact on the characteristics of our

Figure 4. Power comparison of \hat{Z}_1 and Z_1 for design 2 at $N = 500$



test. In our simulations, censoring times were generated uniformly across groups (according to the Weibull(2, 5) distribution) and although our censoring hazard model specified group-specific hazards, our test statistic performed adequately. It is likely that a single-sample hazard model – the simple Nelson-Aalen estimator for the censoring hazard – would produce similar results, (i.e.) that our test would remain appropriately sized under a single-sample censoring hazard estimate. There has been little research into the effect of model mis-specification on IPCW statistics, and it could be expected that bias could result – indeed, the proofs in this and the previous chapter tacitly assume the correctness of the specification of censoring hazard model. Additional studies could evaluate the impact of model mis-specification on the size and power of our test statistic, as well as evaluate its

performance under more complex multi-stage networks, although this is beyond the scope of this dissertation.

Simulation Study – Comparison

We performed an additional simulation study comparing \widehat{Z}_1 to a two-sample IPCW Mann-Whitney statistic for waiting times proposed by Fan and Datta [15]. In our notation, the Fan-Datta test statistics can be defined as

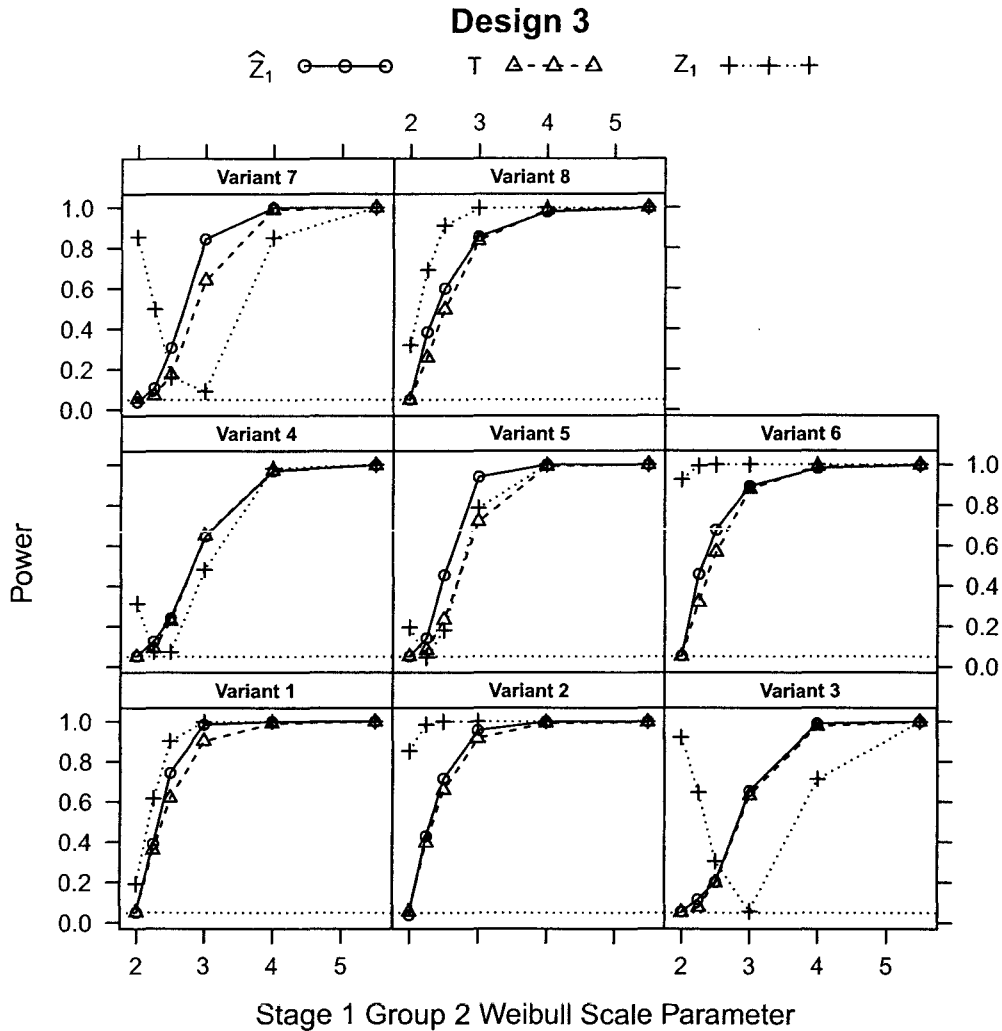
$$\begin{aligned}\widehat{U}_{12j} &= \frac{1}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \frac{I[U_{1i_1j} - T_{1i_1j} \leq U_{2i_2j} - T_{2i_2j}] \delta_{1i_1j} \gamma_{2i_2j}}{\widehat{K}_1(U_{1i_1j}-) \widehat{K}_2((U_{1i_1j} - T_{1i_1j}) + T_{2i_2j})-} \\ \widehat{U}_{21j} &= \frac{1}{n_1 n_2} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \frac{I[U_{2i_2j} - T_{2i_2j} \leq U_{1i_1j} - T_{1i_1j}] \delta_{2i_2j} \gamma_{1i_1j}}{\widehat{K}_2(U_{2i_2j}-) \widehat{K}_1((U_{2i_2j} - T_{2i_2j}) + T_{1i_1j})-}.\end{aligned}\quad (3.12)$$

In (3.12), the \widehat{K}_i refer to Kaplan-Meier estimators of the censoring hazard for group i . Fan and Datta suggest the test statistic $T = 0.5 (\widehat{U}_{12j} + 1 - \widehat{U}_{21j})$ and prove and illustrate via simulation its asymptotic normality. The need for both \widehat{U}_{12j} and \widehat{U}_{21j} stems from the lack of symmetry in U_{12j} caused by right censoring. In general, \widehat{U}_{12j} and $1 - \widehat{U}_{21j}$ will be close but not necessarily equal.

For uncensored data, the log-rank test is optimal (i.e. most powerful) for alternative hypotheses that stipulate proportional hazards (see e.g. Andersen, et al, Sec. V.2.3 [6]). Hence, we expect that our IPCW log-rank test statistic \widehat{Z}_1 will exhibit greater power than the Fan-Datta test statistic T under proportional hazards alternatives. We designed our simulation comparing \widehat{Z}_1 and T to test this expectation. The design for this simulation was identical to the semi-Markov setup of Design 2 from the previous section. Stage 0 waiting times, censoring times, and the correlation between stage 0 and stage 1 waiting times were generated as in Design 2 from the previous section. In this simulation, stage 1 waiting times, the quantity being tested, were generated from the Weibull distribution with shape parameter fixed at 2. The scale parameter was 2 for group 1 and ranged from 2 to 5.5 for group 2, covering the null (scale = 2) and alternative (scale > 2) hypotheses. Note that for a fixed common shape parameter, Weibull hazards are proportional with hazard ratio $(\lambda_1/\lambda_2)^\gamma$, where λ_i denotes the scale parameter in group i and γ the common shape parameter. The per-group sample size for our simulated data was 100.

The results of these simulations are depicted in Figure 5. The empirical power

Figure 5. Power comparison of \hat{Z}_1 and the Fan-Datta test T



function for the naive log-rank test Z_1 is plotted for reference and to further illustrate its inappropriateness for dependent waiting times. Both test statistics maintained appropriate size, and for large alternatives (i.e. for Weibull scale parameters 4 and greater), the power of the two tests differed little, with our log-rank test exhibiting only slightly greater power. The power of our log-rank test did, however, ascend more rapidly to the limiting value of 1, and was greater than the power of the Fan-Datta test for moderately-sized alternatives (Weibull scale parameters from 2.25 to 3.0) under several of the scenarios. Hence, for this particular design of proportional hazards alternatives, we can reasonably conclude that our log-rank test was more strongly powered than the Fan-Datta test, and that the gain in power was largest for moderately-sized alternative hypotheses.

Analysis of Burn Patient Data

To demonstrate the practical use of our test statistic, we analyzed data from a cohort study of 154 burn patients [23]. These data are available online [26, 27] and are described in full detail elsewhere [23, 25]. We provide a brief description here in the context of a multi-stage model. Patients entered the study at time 0 in stage 1 for the treatment of their burns. Subsequently, patients received one of two interventions – excision (stage 2) or prophylactic antibiotic treatment (stage 3) – or developed staphylococcus infection (stage 4). After excision, (i.e.) entry into stage 2, patients either received prophylactic antibiotic and entered stage 5 or developed infection and entered stage 6. Following prophylactic antibiotic as a first treatment (entry into stage 3), patients either had their wounds excised and entered stage 5 or developed staphylococcus infection and entered stage 7. Patients developing infection prior to treatment (stage 4) could only have their wounds excised (stage 6). Patients having received both treatments prior to infection (stage 5) could then develop infection and progress to stage 8 and patients developing infection after antibiotic (stage 7) could have their wounds excised and progress to stage 8. Stages 6 and 8 were terminal, as stage 8 exhausts all of the events that could occur and patients in stage 6 had already developed infection and had no need for prophylactic antibiotic treatment. Table 3 describes the stages in the model Figure 6 depicts the network structure.

We assumed that interventions (antibiotic, excision) occurred before staphylococcus infection when the two occurred simultaneously and that interventions could occur simultaneously. Specifically, two patients developed staphylococcus infections at the same time as excision and were treated as having passed through stage 2 with a waiting time of zero, while one patient that developed staphylococcus infection at the same time as

Table 3

Description of stages in multi-stage burn data model.

-
-
1. Onset of burn treatment
 2. Excision of burn as first intervention
 3. Prophylactic antibiotic as first intervention
 4. Staphylococcus infection prior to excision or antibiotic
 5. Excision and antibiotic (prior to infection)
 6. Excision and staphylococcus infection (no antibiotic)
 7. Staphylococcus infection after antibiotic, no excision
 8. Excision, antibiotic, and staphylococcus infection
-

Figure 6. Network of stages for multi-stage burn patient data

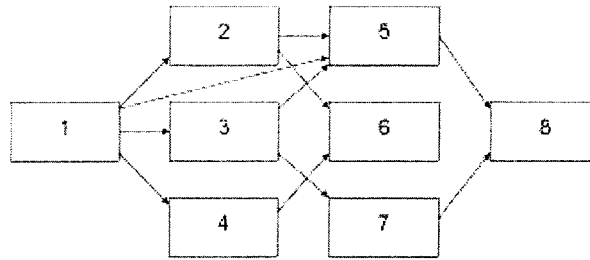


Table 4

Observed transitions of 154 burn patients

From	To							
	1	2	3	4	5	6	7	8
1	24	53	39	28	10	0	0	0
2		33	0	0	14	6	0	0
3			12	0	21	0	6	0
4				17	0	11	0	0
5					37	0	0	8
6						17	0	0
7							2	4
8								12

receiving antibiotic was treated as having passed through stage 3 with a waiting time of zero. Ten patients simultaneously received excision and prophylactic antibiotic and were treated as having passed directly from stage 1 to stage 5. The data set is provided in the Appendix.

Most patients (102 of 154) were treated by excision or antibiotic before infection developed or before they were censored (Table 4). Among those who received excision as a first treatment (stage 2), most were censored (33 of 53), while the majority of patients receiving antibiotic as a first treatment next had their wounds excised (moved to stage 5, 21 of 39). For the 45 patients receiving both excision and prophylactic antibiotic prior to infection, eight developed infection and the remainder were censored. Four of the six patients developing infection secondary to antibiotic treatment received excision. Twenty-nine patients reached terminal stages in the model – 17 developed infection after excision, eight developed infection after both treatments, and 4 received excision after prophylactic antibiotic was unsuccessful in preventing infection.

Our focus was the waiting times for patients entering stage 2, those having received

excision as their first intervention prior to infection. We compared stage 2 waiting times with our test statistic, $\widehat{Z}_2(t)$, among four groups defined by the crossing of two binary factors in the data set: (1) receipt of routine bathing or full body cleansing, which was the intervention in the study and (2) the experience of a respiratory burn. Fifty-three patients entered stage 2 – Table 5 provides the distribution across the 4 groups, the transitions that followed entry into stage 2, and the correlation between stage 1 and stage 2 waiting times. We noted that patients experiencing respiratory burns (groups 2 and 4) were most likely to receive antibiotic as a second intervention (i.e.) move to stage 5 after stage 2 (8 of 16), while those not experiencing respiratory burns (groups 1 and 3) were most likely to be censored after entry into stage 2 (28 of 37). Censoring rates were roughly equivalent for those receiving routine bathing (12/21 = 57%) and those receiving full body cleansing (21/32 = 66%), as were transitions into stages 5 and 6.

Figure 7 depicts stage 2 waiting times and the Satten and Datta [44] estimator $\widehat{S}_2(t)$. Empirically, those experiencing respiratory burns (dashed lines) exited stage 2 more rapidly than those not (solid lines) as shown by the shorter lines in the left panel of Figure 7 and the steeper survival function estimates in the right panel. Differences between those receiving routine bathing (dots) and full body cleansing (triangles) were minimal. The chi-square statistic on 3 degrees of freedom for the naive log-rank test ($Z_2(t)$) was 7.4, indicating marginal significance ($p = .06$). The chi-square statistic associated with $\widehat{Z}_2(t)$, for which bootstrap resampling at 1000 iterations was used to estimate the variance, was 9.4, indicating a significant difference among the 4 groups ($p = .02$). The proximity of the p-values for the naive log-rank test and our test suggests that the effect of potential dependence in censoring or transition times had only a marginal effect on the comparison of stage 2 waiting time distributions. Indeed, the Kaplan-Meier estimates of the survival

Table 5

Disposition of 53 patients observed to enter stage 2 in the burn data model

Group	Description	N	Next Stage			Stage 1
			Stage 5	Stage 6	Cens.	Correlation
1	Bath, No Resp. Burn	12	2	1	9	-0.40
2	Bath, Resp. Burn	9	4	2	3	0.13
3	Cleanse, No Resp. Burn	25	4	2	19	-0.05
4	Cleanse, Resp. Burn	7	4	1	2	-0.20
All		53	14	6	33	-0.16

Figure 7. Stage 2 waiting times and survival functions for 4 groups of burn patients. Plotted points represent patients who were censored. Patients presented with (dashed line) or without (solid line) respiratory burns and received routine bathing (o) or full body cleansing (Δ).

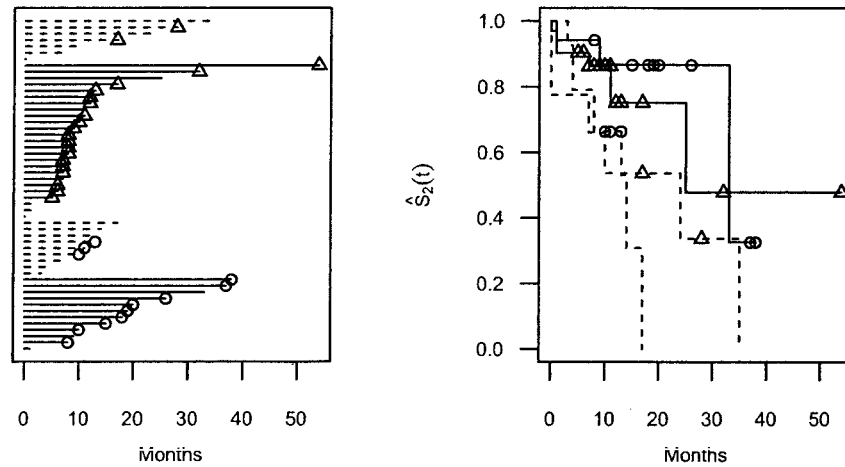


Table 6

Comparison of Kaplan-Meier and Satten-Datta stage 2 survival function estimates

Group 1			Group 2			Group 3			Group 4		
t	$S_2^*(t)$	$\hat{S}_2(t)$	t	$S_2^*(t)$	$\hat{S}_2(t)$	t	$S_2^*(t)$	$\hat{S}_2(t)$	t	$S_2^*(t)$	$\hat{S}_2(t)$
1	0.92	0.94	3	0.89	0.93	0	0.96	0.97	0	0.86	0.78
9	0.83	0.87	4	0.78	0.79	1	0.88	0.90	7	0.71	0.66
33	0.55	0.33	8	0.67	0.66	7	0.83	0.86	10	0.57	0.54
			13	0.50	0.53	11	0.74	0.75	24	0.38	0.34
			14	0.25	0.31	25	0.49	0.48	35	0.00	0.00
			17	0.00	0.00						

function for each group in large part differed only slightly from the values provided by the Satten-Datta estimator (Table 6), indicating that the effect of adjusting for censoring by IPCW was minimal. Note also that although correlations between stage 1 and stage 2 waiting times varied over the groups, the correlations themselves were not particularly strong (Table 5). Additionally, Groups 1 and 4 exhibited the smallest differences in Satten-Datta and Kaplan-Meier estimates of $S_2(t)$ as well as the largest absolute rank correlations, giving some indication that waiting time correlations do have some impact on estimation of survival curves.

In calculating the test statistic $\hat{Z}_2(t)$, we modeled the censoring hazard $K_{hi}(t)$ as a function of 14 time-invariant external covariates – type of bathing solution applied (routine bath, full body cleansing), sex, race (white or non-white), percentage of body surface area

burned, indicator functions for the presence of head, buttock, trunk, upper leg, lower leg, and respiratory tract burns, and a four level factor accounting for the type of burn (chemical, scalding, electrical, flame) – and time-dependent internal covariates denoting stage occupation just before time t . This selection of covariates represents all available covariates provided with this data set. We do note that neither interactions of covariates nor functions of the internal stage occupation covariates were considered in modeling the censoring hazard, mostly to avoid over-specification of the censoring hazard model. In general, consideration of such terms requires some biological or clinical justification and we can consider our specification to be adequate.

CHAPTER IV

NONPARAMETRIC REGRESSION

Three of the more commonly-used regression models used for survival data are parametric models, the semi-parametric Cox proportional hazards model [9], and the accelerated failure time (AFT) model, with the Cox model enjoying considerable popularity. Aalen's nonparametric linear model [2, 3, 4], which we have used in previous chapters as the estimator of the censoring hazard in multi-stage models to generate IPCW estimators and test statistics, has received far less attention. As noted in the introduction, the potential for negative hazard estimates, problems with rank-deficient regression matrices, and the difficulty in interpreting integrated regression coefficients are likely sources of this lack of attention. Nevertheless, as noted in our discussion of the censoring hazard model, Aalen's model provides a very flexible alternative to the more commonly used models. This flexibility stems from the fact that Aalen's model defines and estimates time-varying regression coefficient functions rather than static coefficient estimates, in addition to the allowing for time-varying covariates as the Cox model and versions of the AFT model allow. Further, the ad hoc adjustments to hazard estimates derived from Aalen's linear model noted in the introduction – fixing a lower bound of zero for the estimated hazard and ignoring time points for which the regression matrix is of less than full column rank – along with kernel smoothing techniques for the coefficient functions have been suggested and generally perform well [20].

In this chapter, we define nonparametric regression coefficient estimators for waiting times from a multi-stage model, analogous to the estimators for Aalen's linear model for survival data. As before, we apply the IPCW principle to derive our estimators. We begin by defining Aalen's linear model and describing estimators for waiting times from uncensored multi-stage data. We then define IPCW coefficient estimators, derive their martingale representation, and establish their asymptotic normality. We then conduct a simulation study evaluating the performance of our estimators and demonstrating the inappropriateness of the use of Aalen's linear model for survival data for multi-stage data.

Finally, we demonstrate the use of our estimators through the analysis of a data set of 137 bone marrow transplant patients [8]. Note that we now revert back to the single-sample notation of Chapter II.

Estimators for Uncensored Data

We begin by defining some vector and matrix notation for the counting processes that will form our coefficient estimators. As before, we establish notation and models for uncensored data to provide a foundation for our IPCW estimators. The introduction of Aalen's linear model follows the notation and discussion of Andersen, et al [6], who provide one of the more extensive treatments of Aalen's model. Although the notation of our model for the waiting time hazard will differ from the notation for the censoring hazard model, the formulations are equivalent. Define the individual, uncensored data stage exit and at risk counting processes $N_{ij}^*(t) = I[U_{ij}^* - T_{ij}^* \leq t, X_{ij} = 1]$ and $Y_{ij}^*(t) = I[U_{ij}^* - T_{ij}^* \geq t, X_{ij} = 1]$ for $1 \leq i \leq n$. Note that these individual-level counting processes are the quantities summed to produce the aggregate counting processes from Chapter II, (i.e.) $N_j^*(t) = \sum_i N_{ij}^*(t)$ and $Y_j^*(t) = \sum_i Y_{ij}^*(t)$. Collect the stage j exit counting processes into the vector $\mathbf{N}_j^*(t) = (N_{1j}^*(t), \dots, N_{nj}^*(t))$. Recall that the vector of covariates for the i^{th} individual is $\mathbf{X}_i(t) = (1, X_{i1}(t), \dots, X_{ip}(t))$ and that $\mathbf{X}(t)$ is the $n \times (p+1)$ matrix of covariates for all individuals with i^{th} row $\mathbf{X}_i(t)$. Let $\mathbf{Y}^*(t)$ denote the "at risk" covariate matrix, with i^{th} row $Y_{ij}^*(t) \cdot (1, X_{i1}(T_{ij}^* + t), \dots, X_{ip}(T_{ij}^* + t))$. The i^{th} row of $\mathbf{Y}_j^*(t)$ provides the covariate vector for individual i multiplied by the at risk process for individual i . Note that we have necessarily added T_{ij}^* to the argument of the covariate functions X_{ik} , since the covariate processes evolve in calendar time but their impact on the stage j exit hazard only occurs after stage j entry, (i.e.) after time T_{ij}^* . A matrix definition of $\mathbf{Y}_j^*(t)$ will be useful in developing our results. Abusing notation, we note that

$\mathbf{Y}_j^*(t) = \mathbf{diag}\{(Y_{1j}^*(t), \dots, Y_{nj}^*(t))\} \cdot \mathbf{X}(\mathbf{T}_j^* + t)$, where \mathbf{T}_j^* denotes the vector of true stage j entry times. By this notation, we indicate that the i^{th} row of \mathbf{X} is evaluated at time $T_{ij}^* + t$.

The model for the hazard of stage j exits is defined as a simple linear model, similar to the model for the censoring hazard. At the individual level and in matrix form (again abusing notation), we have:

$$\lambda_{ij}(t) = \beta_{0j}(t) + \sum_{m=1}^p \beta_{mj}(t) X_{im}(T_{ij}^* + t), \quad 1 \leq i \leq n$$

$$\Lambda_j(t) = \mathbf{X}(\mathbf{T}_j^* + t)\beta_j(t),$$

where $\Lambda_j(t) = (\lambda_{1j}(t), \dots, \lambda_{nj}(t))$ and $\beta_j(t) = (\beta_{0j}(t), \dots, \beta_{pj}(t))$. Under the multiplicative intensity model for counting processes (see, e.g. Andersen, et al [6]), we have that

$\mathbf{N}_j^*(t) = \int_0^t \text{diag}\{Y_{1j}^*(s), \dots, Y_{nj}^*(s)\}\Lambda_j(s) + \mathbf{M}_j^*(t)$, where we define

$\mathbf{M}_j^*(t) = (M_{1j}^*(t), \dots, M_{nj}^*(t))$ to be the vector of individual martingales associated with stage j exits. Substituting in the expression for $\Lambda_j(t)$ under Aalen's linear model admits the martingale estimating equation

$$\mathbf{N}_j^*(t) = \int_0^t \mathbf{Y}_j^*(s)d\mathbf{B}_j(s) + \mathbf{M}_j^*(t), \quad (4.1)$$

where $\mathbf{B}_j(t) = (B_{0j}(t), \dots, B_{pj}(t))$ is the vector of integrated regression coefficients

$B_{mj}(t) = \int_0^t \beta_{mj}(s)ds, 0 \leq m \leq p$; in particular, we note that $d\mathbf{B}_j(t) = \beta_j(t)dt$. Setting the "noise" (i.e. the martingale $\mathbf{M}_j^*(t)$) equal to zero in this equation produces the estimator for the integrated regression coefficients for uncensored data,

$$\mathbf{B}_j^*(t) = \int_0^t J_j^*(s)\mathbf{Y}_j^{*-}(s)d\mathbf{N}_j^*(s), \quad (4.2)$$

where $\mathbf{Y}_j^{*-}(t)$ is any generalized inverse of $\mathbf{Y}_j^*(t)$ and $J_j(s) = I[\text{rank}(\mathbf{Y}_j^*(s)) = p + 1]$ is added to avoid complications from regression matrices of less than full column rank. The estimator (4.2) is often referred to as a generalized Nelson-Aalen estimator, given its functional similarity with the Nelson-Aalen estimator. In fact, when the regression matrix $\mathbf{X}(s)$ is a vector of ones, the coefficient estimator (4.2) is the Nelson-Aalen estimator.

By the martingale equation (4.1), we have the martingale representation

$(\mathbf{B}_j^* - \mathbf{B}_j)(t) = \int_0^t J_j(s)\mathbf{Y}_j^{*-}(s)d\mathbf{M}_j^*(s)$, which subsequently provides us with the predictable and optional variation processes for $\mathbf{B}_j^*(t)$,

$$\langle \mathbf{B}_j^* - \mathbf{B}_j \rangle (t) = \int_0^t J_j^*(s)\mathbf{Y}_j^{*-}(s) \text{diag}\{\lambda_j(s)\} \mathbf{Y}_j^{*-}(s)^T ds, \quad (4.3)$$

$$[\mathbf{B}_j^* - \mathbf{B}_j] (t) = \int_0^t J_j^*(s)\mathbf{Y}_j^{*-}(s) \text{diag}\{d\mathbf{N}_j^*(s)\} \mathbf{Y}_j^{*-}(s)^T, \quad (4.4)$$

where $\lambda_j(s) = (\lambda_{1j}(s), \dots, \lambda_{nj}(s))$. Note that the optional variation process (4.4) serves as an estimate of the covariance matrix for $\mathbf{B}_j^*(t)$. Under suitable regularity conditions on the matrix $\mathbf{Y}_j^*(t)$, the estimator $\mathbf{B}_j^*(t)$ approaches a Gaussian limit with covariance matrix given by the expected value of the predictable variation process (4.3). For proof of this

result, we refer the reader to Theorem VII.4.1 of Andersen, et al [6]. We do note that the regularity conditions required of $\mathbf{Y}_j^*(t)$ depend on the choice of generalized inverse, and that Andersen, et al utilize a least squares principle and define

$$\mathbf{Y}_j^-(t) = (\mathbf{Y}_j^*(t)^T \mathbf{Y}_j^*(t))^{-1} \mathbf{Y}_j^*(t)^T \text{ for their proofs.}$$

IPCW Estimators

The estimator $\mathbf{B}_j^*(t)$ has no use for right-censored data and the validity of estimators based on unweighted, censored data counting processes depends on independent censoring and the semi-Markov property. As before, to relax these requirements we propose IPCW estimators for the integrated regression coefficients $\mathbf{B}_j(t)$ from Aalen's model. The IPCW estimators are constructed just as the estimators for the stage j cumulative hazard and survival functions from Chapter II, by replacing the counting processes from the uncensored data estimators with IPCW analogues. Before establishing our estimators, we define the individual IPCW counting processes $\widehat{N}_{ij}(t) = I[U_{ij} - T_{ij} \leq t, \delta_{ij} = 1] / \widehat{K}_i(U_{ij}-)$ for stage j exits and the IPCW stage j at-risk counting processes $\widehat{Y}_{ij}(t) = I[U_{ij} - T_{ij} \geq t, \gamma_{ij} = 1] / \widehat{K}_i(T_{ij} + t-)$ for $1 \leq i \leq n$. Note that the $\widehat{K}_i(t)$ are as defined in Chapter II from Aalen's linear model for the censoring hazard. Define the vector $\widehat{\mathbf{N}}_j(t) = (\widehat{N}_{1j}(t), \dots, \widehat{N}_{nj}(t))$ and the weighted, modified coefficient matrix $\widehat{\mathbf{Y}}_j(t) = \text{diag}\{\widehat{Y}_{1j}(t), \dots, \widehat{Y}_{nj}(t)\} \cdot \mathbf{X}(\mathbf{T}_j + t)$. Our IPCW estimator for the regression coefficient vector is

$$\widehat{\mathbf{B}}_j(t) = \int_0^t \widehat{J}_j(s) \widehat{\mathbf{Y}}_j^-(s) d\widehat{\mathbf{N}}_j(s), \quad (4.5)$$

where $\widehat{\mathbf{Y}}_j^-(t)$ is a generalized inverse of $\widehat{\mathbf{Y}}_j(t)$ and $\widehat{J}_j(s) = I[\text{rank}(\widehat{\mathbf{Y}}_j(s)) = p + 1]$.

The asymptotic properties of the IPCW estimators of Chapter II and test statistics in Chapter III relied on martingale representations and the martingale central limit theorem. Hence, to explore the asymptotic properties of $\widehat{\mathbf{B}}_j(t)$, we establish its martingale representation in the following lemma.

Lemma IV.1 The estimator $\widehat{\mathbf{B}}_j(t)$ defined in (4.5) has martingale representation

$$\begin{aligned} \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) &= \int_0^t \widehat{J}_j(s) \mathbf{Y}_j^{*-}(s) d\mathbf{M}_j^*(s) + \\ &\int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))^{-1} \text{diag}\{\mathbf{Q}(u, s)\} \text{diag}\{d\mathbf{N}_j^*(s)\} \right] \times \\ &\mathbf{P}(u) d\mathbf{M}^c(u) + o_p(n^{-1/2}), \end{aligned}$$

where $\mathbf{D}_Y(u, s) = \int_0^\infty \mathbf{diag} \{ I[u < T_{ij}^* + s] / K_i(u) dM_i^c(u) \}$, $\mathbf{Q}(u, s)$ is the vector with i^{th} element $(I[u < U_{ij}^*] - I[u < T_{ij}^* + s]) / K_i(u)$, $\mathbf{P}(u) = \mathbf{I} - \mathbf{Y}^c(u)(\mathbf{Y}^c(u)^T \mathbf{Y}^c(u))^{-1} \mathbf{Y}^c(u)^T$, and $\mathbf{Y}^c(u)$ is the $N \times (p+1)$ matrix $\mathbf{diag} \{ I[T_i \geq u] \} \mathbf{X}(u)$.

The proof of Lemma IV.1 is a ‘‘vectorized’’ version of the proof of the martingale representation of the Satten-Datta estimator $\widehat{\Lambda}_j(t)$ in Lemma II.4 and deferred to the Appendix. As before, we note that this representation decomposes $\widehat{\mathbf{B}}_j(t)$ into the sum of two orthogonal martingales, one related to the counting process for stage j exits (\mathbf{M}_j^*) and one for the counting process for censoring events (\mathbf{M}^c). Further, the first term of the martingale representation for $\widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t)$ is exactly the martingale representation for $\mathbf{B}_j^*(t) - \mathbf{B}_j(t)$, the uncensored data estimator. Due to the orthogonality of the component martingales, when developing the asymptotic properties of $\widehat{\mathbf{B}}_j(t)$ and in particular its predictable and optional variation processes we need not consider the covariation between these two terms. Subsequently, the predictable and optional covariation processes can be defined as

$$\begin{aligned} \langle \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) \rangle &= \int_0^t J_j^*(s) \mathbf{Y}_j^{*-}(s) \mathbf{diag} \{ \boldsymbol{\lambda}_j(s) \} \mathbf{Y}_j^{*-}(s)^T ds + \\ &\int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag} \{ \mathbf{Q}(u, s) \} \mathbf{diag} \{ d\mathbf{N}_j^*(s) \} \right] \times \\ &\mathbf{P}(u) \mathbf{diag} \{ \boldsymbol{\lambda}^c(u) \} \mathbf{P}(u)^T \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag} \{ \mathbf{Q}(u, s) \} \right. \\ &\left. \mathbf{diag} \{ d\mathbf{N}_j^*(s) \} \right]^T du, \end{aligned} \tag{4.6}$$

$$\begin{aligned} [\widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t)] &= \int_0^t \widehat{J}_j(s) \widehat{\mathbf{Y}}_j^-(s) \mathbf{diag} \{ d\widehat{\mathbf{N}}_j(s) \} \widehat{\mathbf{Y}}_j^-(s)^T + \\ &\int_0^\infty \left[\int_0^t \widehat{\mathbf{Y}}_j^-(s) (\mathbf{I} - \widehat{\mathbf{D}}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag} \{ \widehat{\mathbf{Q}}(u, s) \} \mathbf{diag} \{ d\widehat{\mathbf{N}}_j(s) \} \right] \times \\ &\mathbf{P}(u) \mathbf{diag} \{ d\mathbf{N}^c(u) \} \mathbf{P}(u)^T \left[\int_0^t \widehat{\mathbf{Y}}_j^-(s) (\mathbf{I} - \widehat{\mathbf{D}}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag} \{ \widehat{\mathbf{Q}}(u, s) \} \right. \\ &\left. \mathbf{diag} \{ d\widehat{\mathbf{N}}_j(s) \} \right]^T, \end{aligned} \tag{4.7}$$

where $\widehat{\mathbf{D}}_Y(u, s) = \int_0^\infty \mathbf{diag} \{ I[u < T_{ij} + s] / \widehat{K}_i(u) d\widehat{M}_i^c(u) \}$ and $\widehat{\mathbf{Q}}(u, s)$ is the vector with i^{th} element $(I[u < U_{ij}] - I[u < T_{ij} + s]) / \widehat{K}_i(u)$. Again, we note that the optional variation process (4.7) serves as an estimator of the variance of $\widehat{\mathbf{B}}_j(t)$, which is the in-probability limit of (4.6).

The martingale representation established in Lemma IV.1 provides an avenue for establishing the asymptotic normality of the IPCW integrated coefficient estimator $\widehat{\mathbf{B}}_j(t)$ via the martingale central limit theorem. Define the in-probability limit of the predictable variation process (4.6): $\Sigma_{\mathbf{B}}(t) = p \lim_{n \rightarrow \infty} \langle \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) \rangle$. We then have the following theorem:

Theorem IV.2 *Under suitable regularity conditions,*

$$\sqrt{n}(\widehat{\mathbf{B}}_j - \mathbf{B}_j)(t) \xrightarrow{\mathcal{D}} \mathbf{U}(t),$$

where $\mathbf{U}(t) = (U_0(t), \dots, U_p(t))$ is a $(p+1)$ -variate Gaussian martingale with covariance matrix $\Sigma_{\mathbf{B}}(t)$.

The proof of Theorem IV.2 is provided in the Appendix, along with a statement of the required regularity conditions. We again note that the conditions for the martingale central limit theorem are the convergence of the predictable variation process to a deterministic limit and the Lindeberg condition in which the martingale approaches a continuous limit. For the proof of Theorem IV.2, these conditions largely focus on the regression matrix $\mathbf{Y}_j^*(t)$ and the selected form of its generalized inverse.

Simulation Study

We conducted simulation studies to evaluate the validity of the IPCW coefficient estimator $\widehat{\mathbf{B}}_j(t)$. As in the simulation study of the waiting time log-rank statistic in Chapter III, we considered a three-stage tracking model with individuals entering root stage 0 and proceeded through transient stage 1 to absorbing stage 2 or were censored. We simulated data for 100 individuals under several scenarios to evaluate the performance of $\widehat{\mathbf{B}}_1(t)$ – the stage 1 coefficient estimator – for discrete and continuous predictors, both fixed and time-varying. Under each scenario, we calculated the estimate $\widehat{\mathbf{B}}_1(t)$ and the naive Aalen model estimator $\widetilde{\mathbf{B}}_1(t)$, defined as $\widetilde{\mathbf{B}}_1(t) = \int_0^t J_1(s) \mathbf{Y}_1^-(s) d\mathbf{N}_1(s)$, where $\mathbf{N}_1(s)$ is the vector of individual level censored data counting processes for stage 1 exits, $N_{i1}(s) = I[U_{i1} - T_{i1} \leq s, \delta_{i1} = 1]$, $\mathbf{Y}_1(s)$ is the $n \times (p+1)$ matrix with i^{th} row $Y_{i1}(s)(1, X_{i1}(s), \dots, X_{ip}(s))$ where $Y_{i1}(s)$ is the stage 1 at risk process for censored data given by $Y_{i1}(s) = I[U_{i1} - T_{i1} \geq s, \gamma_{i1} = 1]$, and $J_1(s) = I[\text{rank}(\mathbf{Y}_1(s)) = p+1]$ is the indicator function for the matrix $\mathbf{Y}_1(s)$ having full column rank. We additionally

calculated the variance of $\widehat{\mathbf{B}}(t)$ via the formula for the optional variation process defined in (4.7) in order to construct asymptotic pointwise confidence intervals. We did the same for the naive estimator $\widetilde{\mathbf{B}}(t)$, noting that the optional variation process for this estimator is $\int_0^t J_1(s) \mathbf{Y}_1^-(s) \text{diag}\{d\mathbf{N}_1(s)\} \mathbf{Y}_1^-(s)^T$. The three simulation scenarios are described below.

1. *Discrete, fixed covariate.* For this design, we simulated data for 4 groups of individuals, each group of size 25 and with varying stage 1 hazard. We simulated a bivariate normal pair with marginal variance equal to 1 and covariance equal to -0.5, 0, and 0.5. The mean parameter for the first coordinate in the pair was set equal to 1, while the mean parameter for the second coordinate was set to $k/4$ for group k , $k = 1, \dots, 4$. These pairs were exponentiated to produce correlated lognormal waiting times for stages 0 and 1. The stage 0 and stage 1 correlations were common to all groups, so there were three variants of this design. Censoring times were generated from the Weibull distribution with shape parameter 2 and scale parameter 6 independently of the stage waiting times. For this design, approximately 32% of individuals were censored before exiting stage 0 and 21%, 25%, and 29% were censored while in stage 1 for waiting time correlations -0.5, 0, and 0.5, respectively. A “cell means” version of the linear hazard model under this design is

$$\lambda_{i1}(t) = \beta_{11}(t)X_{i1} + \beta_{21}(t)X_{i2} + \beta_{31}(t)X_{i3} + \beta_{41}(t)X_{i4}, \quad (4.8)$$

where the X_{ik} are time-invariant indicators denoting membership in group k and the $\beta_{k1}(t)$ are the lognormal hazard functions for group k which take the form $\phi(\ln t - k/4)/\Phi(\ln t - k/4)$, where ϕ and Φ represent the density and distribution functions of the standard normal distribution.

2. *Discrete, time-varying covariate.* This design was identical to design 1, with the exception that we introduced the possibility of a group switch after stage 1 entry. A random group-switch time was generated for each individual, and at the switch time, an individual in group 1 switched to group 4 and vice versa, and an individual in group 2 switched to group 3 and vice versa. We generated the data for this design via the following procedure.

- Generate the pair (L_0, L_1) from the bivariate normal distribution with mean vector $(1, \mu_k)$, where μ_k is the log-mean for group k , marginal variances equal to

1, and correlation equal to ρ , where ρ is -0.5, 0, or 0.5 depending on the design variant being run. Exponentiate (L_0, L_1) to produce “baseline” stage 0 and 1 waiting times denoted (W_0, W_1) .

- Generate the group-switch times S from the $U(0, 5)$ distribution.
- For individuals who switched before entering absorbing stage 2 ($S < W_1$), generate the pair (L_0^*, L_1^*) from the bivariate normal distribution with mean vector $(1, \mu_k)$, where μ_k is the log-mean *for the group to which the individual has switched*, marginal variances equal to 1, and correlation equal to ρ , where ρ is -0.5, 0, or 0.5 depending on the design variant being run. Exponentiate (L_0^*, L_1^*) to produce (W_0^*, W_1^*) . If the new stage 1 waiting time exceeds the switching time ($W_1^* > S$), then W_1^* replaces W_1 as the stage 1 waiting time. Otherwise, regenerate (L_0^*, L_1^*) under the appropriate distribution until $W_1^* > S$ and replace W_1 with W_1^* .
- Generate the right censoring times from Weibull distribution with shape parameter 2 and scale parameter 6.

The cell means version of the linear hazard model under this design is similar to design 1:

$$\lambda_{i1}(t) = \beta_{11}(t)X_{i1}(T_{i1} + t) + \beta_{21}(t)X_{i2}(T_{i1} + t) + \beta_{31}(t)X_{i3}(T_{i1} + t) + \beta_{41}(t)X_{i4}(T_{i1} + t), \quad (4.9)$$

where in this design the $X_{ik}(t)$ are time-varying indicators of membership in group k . The $\beta_{k1}(t)$ are the same lognormal hazard functions as in (4.8), and there were 3 variants of this design based on the 3 correlation values for stage 0 and 1 waiting times. Censoring rates under this design were the same as under design 1.

3. *Continuous, fixed covariate.* For this design, we simulated a simple regression model with fixed, continuous covariate. The linear hazard model for these simulations was

$$\lambda_{i1}(t) = \beta_{01}(t) + \beta_{11}(t)X_{i1}, \quad (4.10)$$

where $\beta_{01}(t)$ was the intercept function, $\beta_{11}(t)$ the “slope function”, and X_{i1} the fixed, continuous covariate. To generate these data, we simulated the pair (L_0, L_1) from the normal distribution with zero marginal means, unit variance, and correlation

-0.5, 0, and 0.5, as in designs 1 and 2. We then calculated (U_0, U_1) from the marginal distributions of the L_i , $U_i = \Phi_i(L_i)$, where Φ is the standard normal distribution. We then calculated the stage 0 and 1 waiting times as $F_i^{-1}(U_i)$, where F_i^{-1} represents the quantile function for the exponential distribution. For stage 0 waiting times, F_0^{-1} was the exponential quantile function with rate parameter 0.25. For stage 1 waiting times, F_1^{-1} was the exponential quantile function with rate parameter $0.25 + \beta_0 X_i$, where β_0 took the value 0.05 and X_i represented the covariate value for patient i , generated from the uniform distribution on the interval $(-2, 2)$. Censoring times were generated from the Weibull distribution with shape parameter 2 and scale parameter 6. Under this design, 33% of all individuals were censored before stage 1 entry, and 36%, 32%, and 27% of individuals were censored while in stage 1 for stage waiting time correlations -0.5, 0, and 0.5, respectively. The function $\beta_{01}(t)$ was the constant function 0.25 (the exponential hazard rate) and the function $\beta_{11}(t)$ the constant function 0.05.

Under each design, censoring times were generated independently of stage waiting times and uniformly for all individuals. Hence, we used the Kaplan-Meier estimator to estimate the probability of censoring $K_i(t)$ for the inverse probability of censoring weights. One thousand Monte Carlo iterations were run for each variant of each design.

Before conducting a detailed examination of the performance of the IPCW and naive coefficient estimators, we briefly take an empirical look at the raw results of our simulations. Figures 8 and 9 depict the IPCW and naive estimates for selected regression coefficients and variants of designs 1 through 3 (row 1 through 3 respectively). Variants in which stage 0 and 1 waiting times were negatively correlated are plotted in the left column and variants with positive correlation in the right column. The selected regression coefficients for each design are noted in the y-axis label of each panel of each plot. The IPCW estimator appeared approximately in all situations, but did exhibit greater variance than the naive estimator as illustrated by the greater scatter in the integrated coefficient estimators. The naive estimator seemed to underestimate the true integrated regression coefficient functions for variants in which stage waiting times were negatively correlated – most of the estimated integrated regression functions were below the true values. Conversely, the naive estimator overestimated the true regression coefficients when stage waiting times were positively correlated. We do not show plots from variants in which stage

Table 7

Average bias for IPCW and naive coefficient estimators, simulation designs 1 and 2

Design-Variant	Corr.	t	IPCW Estimator				Naive Estimator			
			$\widehat{\beta}_{11}(t)$	$\widehat{\beta}_{21}(t)$	$\widehat{\beta}_{31}(t)$	$\widehat{\beta}_{41}(t)$	$\widetilde{\beta}_{11}(t)$	$\widetilde{\beta}_{21}(t)$	$\widetilde{\beta}_{31}(t)$	$\widetilde{\beta}_{41}(t)$
1-1	-0.5	2	0.037	0.012	-0.007	0.008	-0.244	-0.209	-0.198	-0.149
		4	0.061	0.072	-0.059	0.063	-0.247	-0.191	-0.213	-0.176
1-2	0	2	0.019	0.022	0.015	0.003	0.009	0.011	0.000	0.000
		4	0.010	-0.007	-0.047	-0.018	-0.009	-0.028	-0.046	-0.009
1-3	0.5	2	0.033	0.020	-0.008	-0.023	0.325	0.280	0.215	0.151
		4	0.033	-0.026	-0.031	-0.054	0.187	0.170	0.129	0.091
2-1	-0.5	2	-0.064	0.008	0.008	0.048	-0.270	-0.187	-0.134	-0.091
		4	-0.029	-0.017	-0.062	0.054	-0.534	-0.393	-0.240	-0.138
2-2	0	2	-0.069	0.010	0.012	0.023	-0.066	-0.004	0.010	0.035
		4	-0.039	-0.020	0.010	-0.037	-0.396	-0.231	-0.103	-0.040
2-3	0.5	2	-0.044	-0.024	0.003	-0.028	0.149	0.190	0.185	0.179
		4	-0.052	-0.033	-0.019	0.011	-0.257	-0.101	0.024	0.064

waiting times were uncorrelated, but do note that both the IPCW and naive estimators appeared approximately unbiased.

To further investigate the bias and variance associated with each estimator, we selected interim time points at which to calculate the integrated coefficient estimates and associated asymptotic 95% confidence intervals, times $t = 2$ and $t = 4$. In Table 7, we present the average bias of our IPCW coefficient estimator and the naive coefficient estimator for each variant of simulation designs 1 and 2. The IPCW estimator was approximately unbiased for each variant – the biases for $\widehat{\beta}_1(t)$ listed in Table 7 represent less than 6% of the true values at the given time points. While overestimation was more likely at time $t = 2$ (14 of 24 estimates) and underestimation at time $t = 4$ (14 of 24) for the IPCW estimator, this pattern was not consistent over groups or variants. The naive estimator was approximately unbiased for variant 2 of designs 1 and variant 2 of design 2 at time $t = 2$, for each of which the waiting time correlation parameter was 0. The naive estimator underestimated the true hazard at time $t = 4$ for variant 2 of design 2, where stage waiting times were uncorrelated.

Figure 8. IPCW coefficient estimators for selected variants of simulation designs 1-3. The left/right column plots variants with negatively/positively correlated waiting times (denoted by -/+). Design 1 estimates are plotted in row 1, design 2 in row 2, and design 3 in row 3. Sixty Monte Carlo iterates are plotted in each, and regression coefficient estimator is identified by the y-axis label. True cumulative hazard estimates for each model parameter are plotted with a solid line.

IPCW Estimator

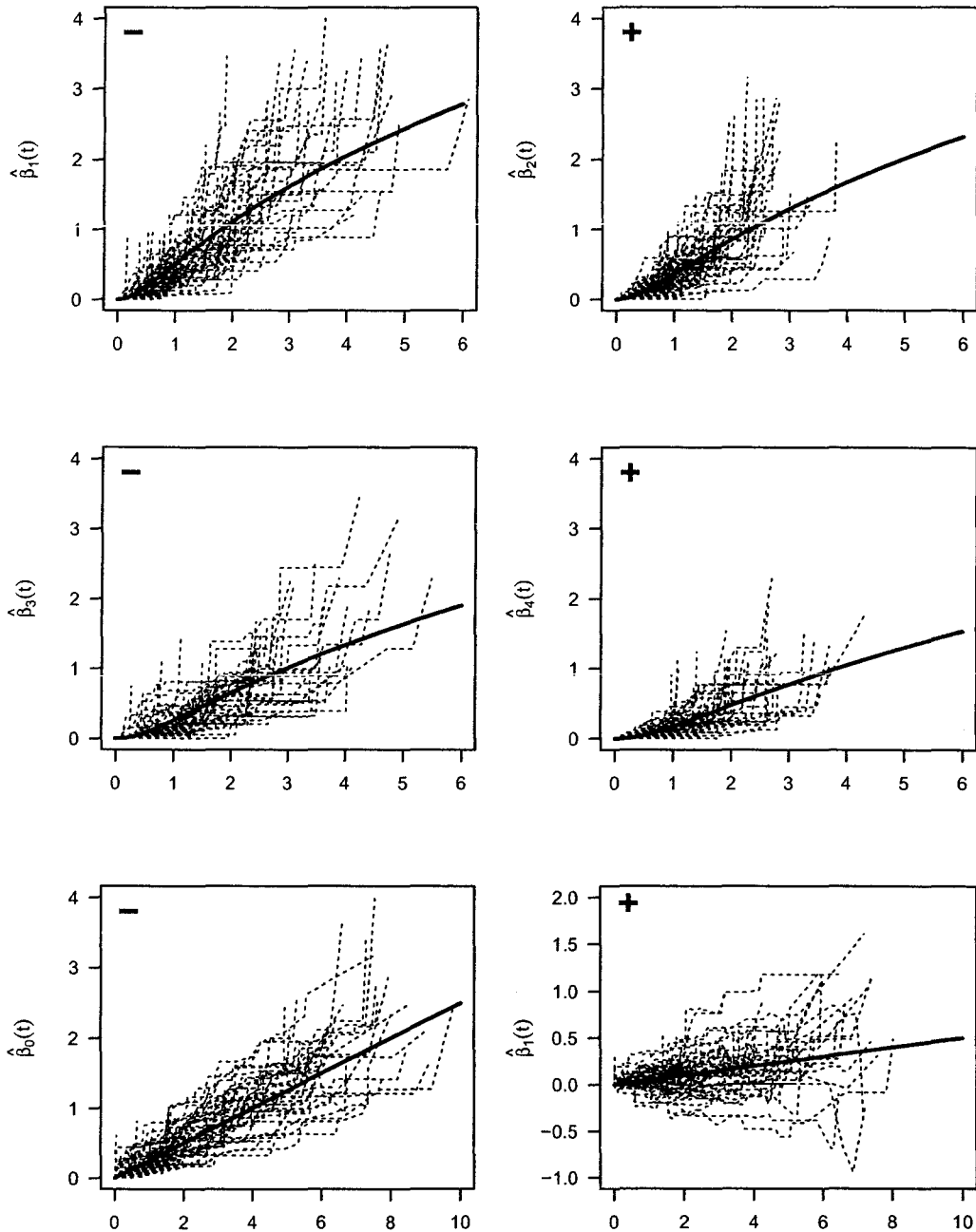


Figure 9. Naive coefficient estimators for selected variants of simulation designs 1-3. The left/right column plots variants with negatively/positively correlated waiting times (denoted by -/+). Design 1 estimates are plotted in row 1, design 2 in row 2, and design 3 in row 3. Sixty Monte Carlo iterates are plotted in each, and regression coefficient estimator is identified by the y-axis label. True cumulative hazard estimates for each model parameter are plotted with a solid line.

Naive Estimator

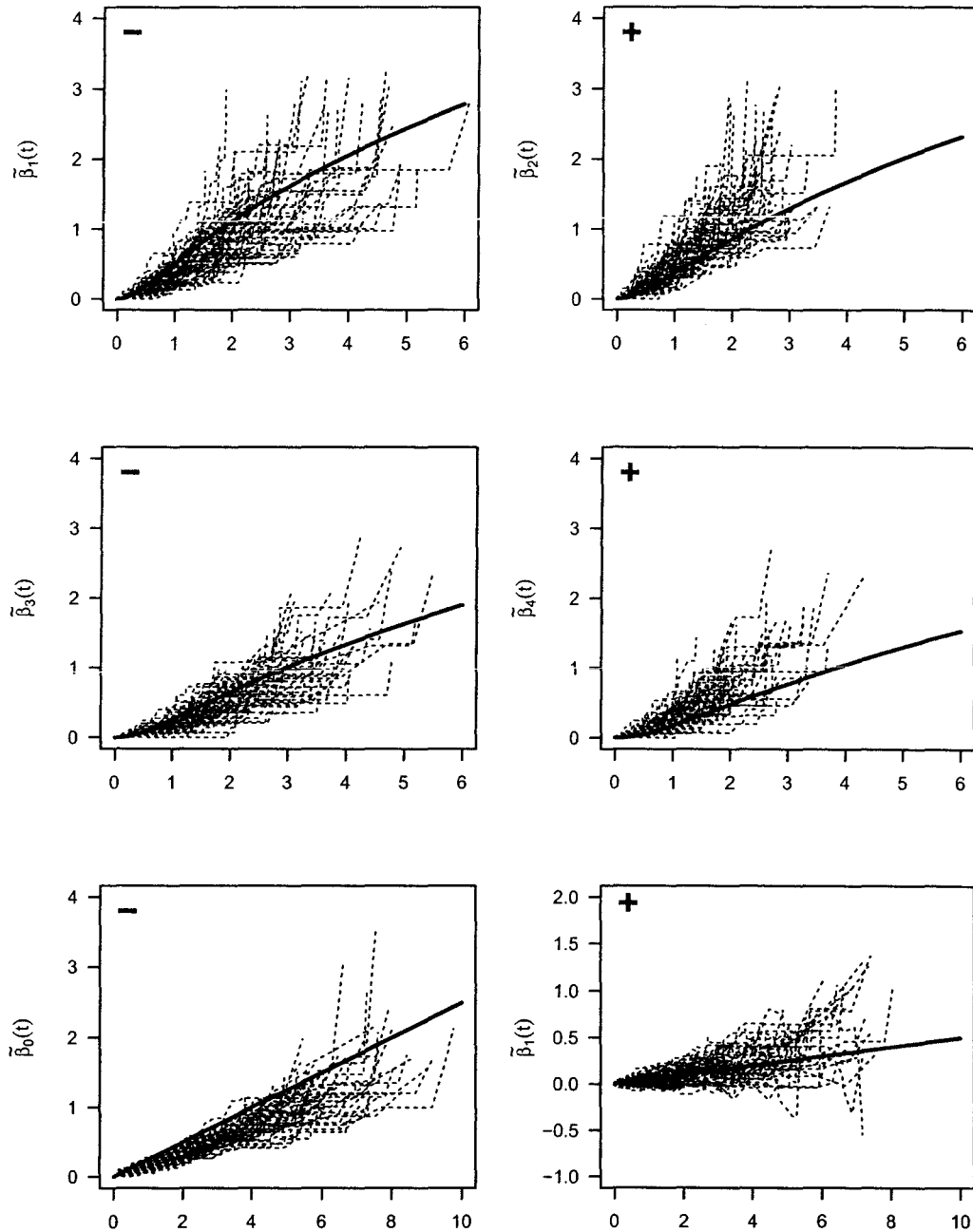


Table 8

Coverage probabilities for asymptotic 95% CI, simulation designs 1 and 2

Design-Variant	Corr.	t	IPCW Estimator				Naive Estimator			
			$\widehat{B}_{11}(t)$	$\widehat{B}_{21}(t)$	$\widehat{B}_{31}(t)$	$\widehat{B}_{41}(t)$	$\widetilde{B}_{11}(t)$	$\widetilde{B}_{21}(t)$	$\widetilde{B}_{31}(t)$	$\widetilde{B}_{41}(t)$
1-1	-0.5	2	0.955	0.960	0.960	0.951	0.922	0.888	0.832	0.863
		4	0.947	0.950	0.954	0.953	0.796	0.813	0.837	0.831
1-2	0	2	0.961	0.967	0.961	0.951	0.954	0.960	0.957	0.959
		4	0.952	0.948	0.955	0.956	0.950	0.954	0.956	0.970
1-3	0.5	2	0.953	0.957	0.961	0.956	0.913	0.909	0.914	0.910
		4	0.967	0.966	0.962	0.960	0.950	0.880	0.901	0.873
2-1	-0.5	2	0.961	0.955	0.954	0.944	0.850	0.904	0.934	0.947
		4	0.945	0.954	0.966	0.953	0.852	0.887	0.937	0.966
2-2	0	2	0.962	0.951	0.950	0.946	0.950	0.959	0.962	0.955
		4	0.933	0.954	0.963	0.960	0.923	0.919	0.913	0.906
2-3	0.5	2	0.965	0.954	0.951	0.946	0.940	0.921	0.917	0.908
		4	0.947	0.953	0.963	0.961	0.849	0.850	0.946	0.946

The naive estimator additionally exhibited substantial bias when stage 0 and 1 waiting times were correlated – $\widetilde{\beta}_1(t)$ underestimated the true cumulative hazard for negatively correlated waiting times and overestimated for positively correlated waiting times. This bias pattern was reasonable for our designs. When waiting times were negatively correlated, longer stage 1 waiting times – which signified lower hazard – were less likely to be censored since they were associated with shorter stage 0 waiting times. Conversely, shorter stage 1 waiting times – which signified greater hazard – were more likely to be censored, being associated with longer stage 0 waiting times. These effects led to the underestimation of the stage 1 cumulative hazard function. An opposite effect was seen for positively correlated stage waiting times – shorter stage 1 waiting times (greater hazard) were more likely to be censored being associated with shorter stage 0 waiting times and vice versa leading to overestimation of the stage 1 cumulative hazard. The presence of a time-varying covariate did not have a substantial impact on the bias exhibited by the IPCW and naive estimators – there were no consistent patterns in bias when comparing designs 1 and 2.

Table 9

Bias and coverage probabilities for asymptotic 95% CI, simulation design 3

Design-Variant	Corr.	t	Bias				Coverage Probability			
			IPCW		Naive		IPCW		Naive	
			$\hat{B}_{11}(t)$	$\hat{B}_{21}(t)$	$\tilde{B}_{11}(t)$	$\tilde{B}_{21}(t)$	$\hat{B}_{11}(t)$	$\hat{B}_{21}(t)$	$\tilde{B}_{11}(t)$	$\tilde{B}_{21}(t)$
3-1	-0.5	2	0.032	0.007	-0.160	-0.021	0.937	0.936	0.476	0.913
		4	-0.030	0.006	-0.259	-0.024	0.938	0.952	0.623	0.917
3-2	0	2	0.006	0.001	0.000	0.002	0.953	0.944	0.944	0.947
		4	0.011	0.008	-0.006	0.009	0.958	0.954	0.957	0.957
3-3	0.5	2	-0.028	0.001	0.180	0.029	0.965	0.942	0.730	0.892
		4	0.008	-0.009	0.401	0.100	0.974	0.968	0.998	0.997

Table 8 provides coverage probabilities for the nominal asymptotic 95% confidence intervals for the IPCW and naive estimators. Confidence intervals for the IPCW estimator were close to the nominal 95% level and, if anything, were slightly overly conservative. Coverage probabilities were fairly consistent with no apparent patterns over the designs, variants, or selected time points. Confidence intervals for the naive estimator were at the nominal level for variant 2 of designs 1 and 2 (no correlation between stage 0 and 1 waiting times), but provided generally did not provide nominal coverage when waiting times were correlated. This aberrant coverage generally can be attributed to the bias associated with naive estimator. Coverage of the naive confidence intervals was generally poorest when the true hazard was underestimated. The coverage bias was lower for overestimated hazards, partly because the variance of the naive estimator is a function of the cumulative hazard itself. Hence, when hazards were overestimated, variances tended to be overestimated, partially (but not completely) correcting the bias in the confidence intervals. Again, the presence of a time-varying covariate seemed to have no impact on the observed coverage of the naive and IPCW confidence intervals, as no patterns emerged when comparing coverage rates for designs 1 and 2.

Table 9 provides the bias and coverage probabilities for the IPCW and naive estimators for variants of design 3. The IPCW was approximately unbiased for all variants of design 3 – the bias estimates in Table 9 represent less than 8% of the true coefficient estimates. The coverage probabilities associated with the asymptotic confidence intervals for the IPCW estimator were close to the nominal 95% level. The intervals tended to

exhibit less than nominal coverage for variant 1, where waiting times were negatively correlated, and greater than nominal coverage for variant 3, where waiting times were positively correlated. The naive estimator was unbiased and its asymptotic confidence intervals close to the nominal level only when waiting times were uncorrelated (variant 2). As noted in the empirical examination (Figure 9), the naive estimator underestimated the true integrated regression coefficient function when stage waiting times were negatively correlated (variant 1), and overestimated the true integrated regression coefficient function when stage waiting times were positively correlated (variant 3). The bias estimates in Table 9 were as high as 50% of the true values for the naive estimator. The coverage exhibited by the asymptotic confidence intervals for the naive estimator was poor. Interestingly, at time $t = 4$ for variant 3 the naive confidence intervals exhibited over-coverage, with probabilities approaching 1. We again note that the asymptotic variance of the naive estimator is a function of the estimated hazard – overestimated hazards are in turn accompanied by overestimated variances, which partially explains the overcoverage of these intervals.

Analysis of Bone Marrow Transplant Data

To demonstrate the use of our regression coefficient estimator for waiting times, we analyzed data from an intervention study 137 bone marrow transplant with acute myelocytic leukemia (AML) and acute lymphoblastic leukemia (ALL) [8]. All patients in the study were treated with a regimen of busulfan 4 mg/kg over 4 days and cyclophosphamide 60mg/kg over 2 days followed by bone marrow transplantation from a sibling. Following transplantation, patients were tracked and the time to five events were noted – return of platelet counts to a sustainable level, onset of acute graft-versus-host-disease (GVHD), onset of chronic GVHD, relapse of AML or ALL, and death. Data for additional prognostic indicators and treatment characteristics were collected – patient age, sex, and cytomegalovirus (CMV) status (positive or negative), donor age, sex, and CMV status, waiting time to transplant, the French-American-British (FAB) classification of morphological status, treating hospital, and administration of a prophylactic combining methotrexate (MTX) with cyclosporin and (possibly) methylprednisolone. These data are available online [26, 27] and are described in further detail elsewhere [8, 25].

The stages in the multi-stage model are described in Table 10 and depicted in Figure

Table 10

Description of stages in multi-stage bone marrow transplant model.

-
-
1. Bone marrow transplantation
 2. Acute GVHD as first event
 3. Platelet recovery as first event
 4. Platelet recovery secondary to acute GVHD
 5. Acute GVHD secondary to platelet recovery
 6. Chronic GVHD
 7. Death or relapse of AML/ALL
-

10. After transplantation (stage 1, time 0), patients either developed acute GVHD (stage 2), exhibited platelet recovery (stage 3), developed chronic GVHD (stage 6), or relapsed or died (stage 7). Patients initially developing acute GVHD (stage 2) could then experience platelet recovery (stage 4), develop chronic GVHD, or relapse or die in any sequence allowed by the model. Similarly, patients initially exhibiting platelet recovery (stage 3) could then develop acute GVHD (stage 5), develop chronic GVHD, or relapse or die in any sequence allowed by the model. Note that we have defined the multi-stage model in this fashion, with two orderings of platelet recovery and acute GVHD, so that the defined network was acyclic. More extensive models are possible by considering distinct orderings of these platelet recovery, acute GVHD, and chronic GVHD, but paths under such orderings would exhibit sparse sample sizes. Further, we note that follow-up continued for patients that relapsed prior to death and times to the other events (platelet recovery, acute GVHD, chronic GVHD) were tracked. However, we treated death or relapse as an absorbing stage and gave no further consideration of events occurring after death or relapse.

The observed transitions for these data are provided in Table 11. The vast majority of patients (117) experienced platelet recovery immediately following transplantation (entered stage 3 from stage 1), and very few (7) experienced acute GVHD initially. Overall, platelet levels recovered for 120 patients, twenty-six patients developed acute GVHD, and fifty-nine developed chronic GVHD. Eighty-three of the 137 patients (61%) were followed until death or relapse, with twenty patients censored after platelet recovery (in stage 3), two after acute GVHD secondary to platelet recovery (in stage 5), and thirty-two after the onset of chronic GVHD (in stage 6).

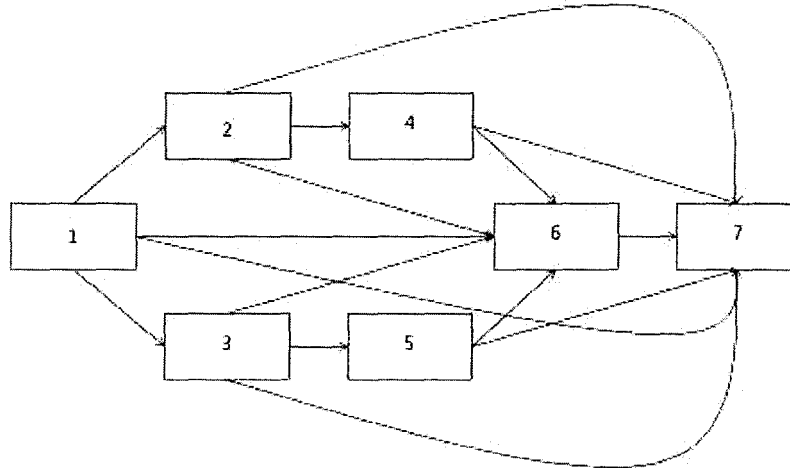
The focus of our analysis was on waiting times in stage 6, the time from onset of

Table 11

Observed transitions of 137 bone marrow transplant patients

From	To						
	1	2	3	4	5	6	7
1	0	7	117	0	0	1	12
2		0	0	3	0	2	2
3			20	0	19	44	34
4				0	0	1	2
5					2	11	6
6						32	27
7							83

Figure 10. Network of stages for multi-stage bone marrow transplant data



chronic GVHD until death or relapse. We modeled stage 6 cumulative hazards as a function of one continuous internal covariate, the time to onset of chronic GVHD after transplantation, one continuous external covariate, patient age, and one three-level discrete external covariate, patient disease group (ALL, low risk AML, high risk AML). Letting λ_{i6} denote the hazard of stage 6 exit for patient i , our model for this analysis was

$$\lambda_{i6}(t) = \beta_{06}(t) + \beta_{16}(t)X_{i1} + \beta_{26}(t)X_{i2} + \beta_{36}(t)X_{i3} + \beta_{46}(t)X_{i4}. \quad (4.11)$$

In this model, we selected the ALL disease group to be the baseline group, and $X_{i1} = T_{i6} - T_{i1}$ represented the time to stage 6 entry, X_{i2} represented patient age at transplantation, X_{i3} was an indicator for low risk AML, and X_{i4} an indicator for high risk

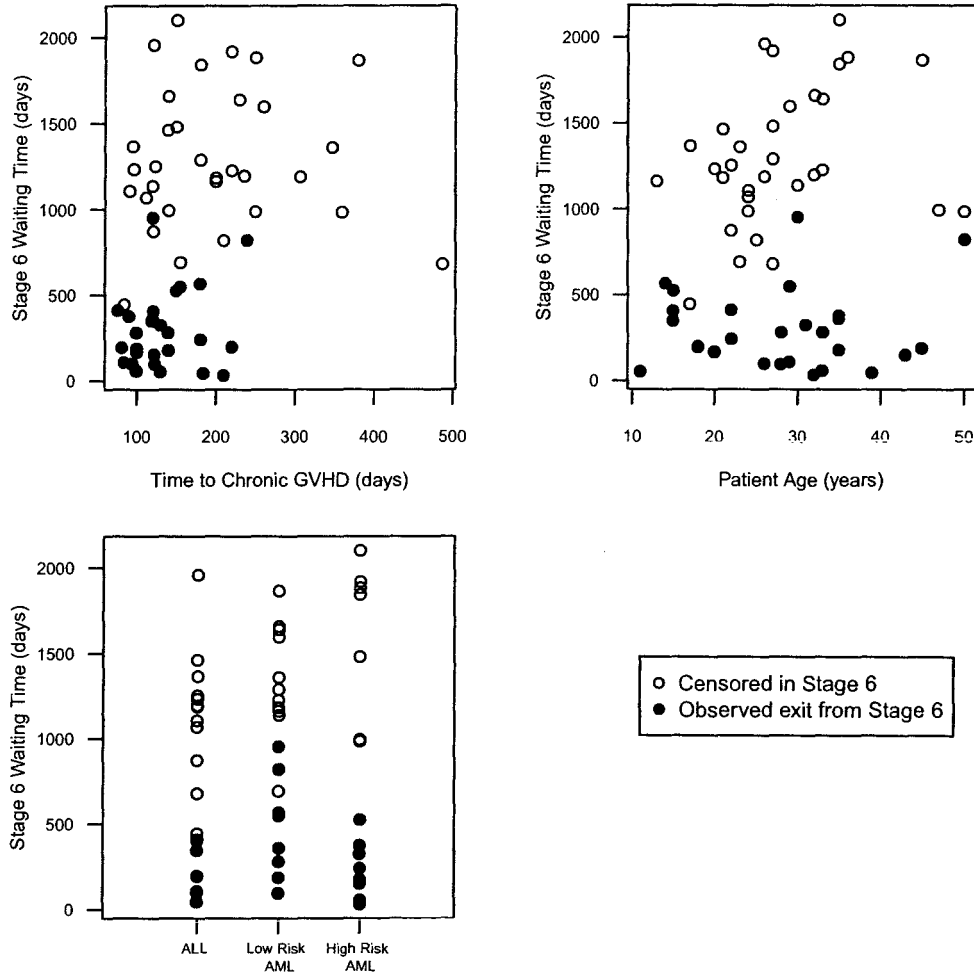
AML. This model provides information on whether patient age and the time to onset of chronic GVHD impacted the risk of death or relapse after chronic GVHD onset, and whether differences in death or relapse hazard after chronic GVHD hazard were observed between disease groups. Our motivation for this specification was to include a variety of types of covariates – continuous and discrete, internal and external.

To construct the inverse probability of censoring weights necessary for our analyses, we modeled the censoring hazard via Aalen’s linear model as a function of 14 fixed external covariates mentioned above – patient disease group (ALL, low risk AML, high risk AML), patient age, sex, and cytomegalovirus (CMV) status (positive or negative), donor age, sex, and CMV status, waiting time to transplant, an indicator function for the FAB classification of morphological status (1 if FAB grade was 4 or 5 and the patient had \overline{AML}), a categorical factor for the treating hospital, and an indicator representing administration of a prophylactic combining methotrexate (MTX) with cyclosporin and (possibly) methylprednisolone. Additionally, we included a 6-vector of internal covariates denoting stage occupation at time t . This model specification exhausted all covariates included in the data set. Interactions among the external covariates and extended functions of the internal covariates were not included in the censoring hazard model to avoid model over-specification and for lack of biological justification.

Of the 59 patients developing chronic GVHD, twenty had ALL, eight of whom died or relapsed, 21 had low risk AML (9 deaths or relapses), and 18 had high risk AML (10 deaths or relapses). The average age was 28 years, which ranged from 11 to 50, and the median time to chronic GVHD was 140 days (min = 76, max = 487). Figure 11 provides an empirical look at the data in the context of our model, plotting stage 6 waiting times against each of the covariates in model (4.11). We noted that longer times to chronic GVHD were seemingly associated with longer times to death or relapse after chronic GVHD onset, although the association was not strict and came with substantial variability. There appeared to be no relationship between patient age and the post-chronic GVHD time to death or relapse, and the disease groups did not appear to significantly differ with respect to post-chronic GVHD time to death or relapse.

The results of our analysis for model (4.11) are depicted in Figure 12 and confirmed empirical observation of Figure 11. The intercept term, had little clinical interpretation and is not plotted, representing the cumulative hazard of ALL patients with zero values for the continuous covariates (age and time to chronic GVHD). We first noted that all stage 6

Figure 11. Plots of stage 6 waiting times against covariates included in model (4.11) – time to chronic GVHD onset, patient age, and disease group. Filled points represent observed deaths or relapses and open points represent censored observations.

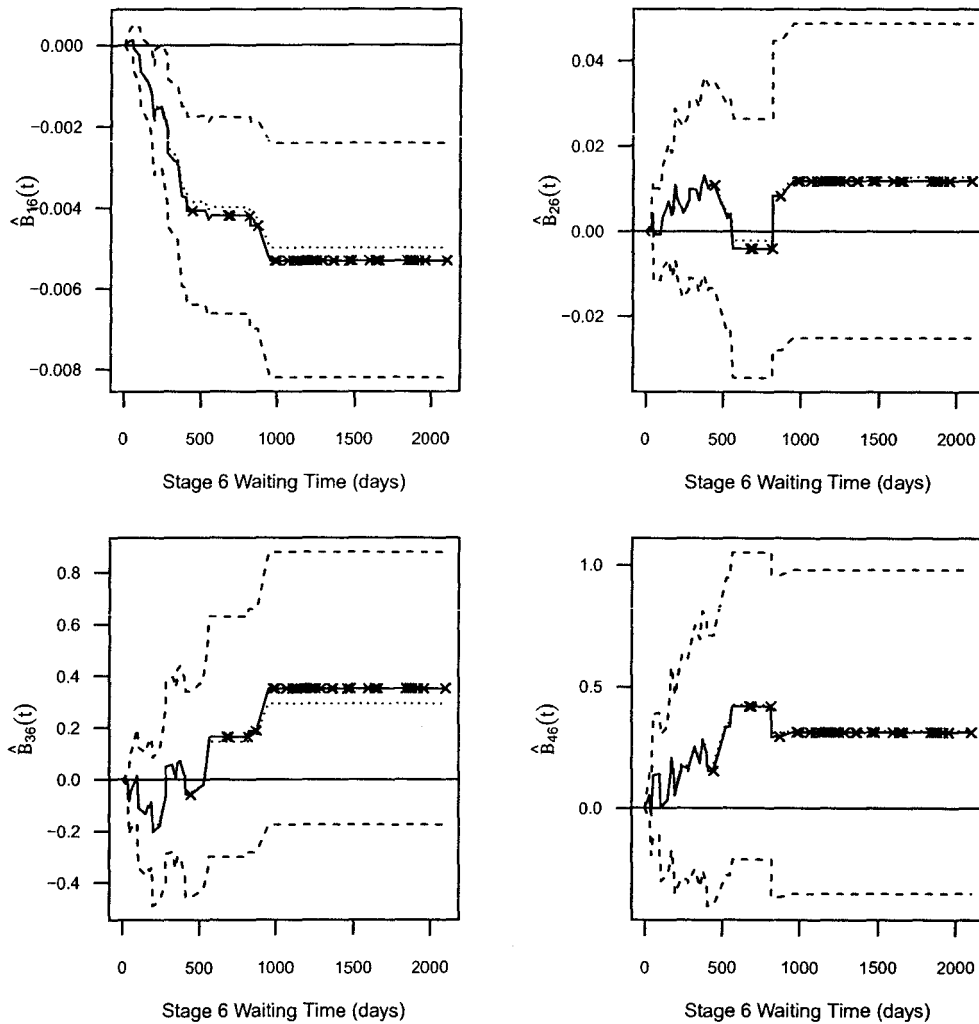


exits occurred within 1000 days upon stage 6 entry and that relatively few observations were censored prior to 1000 days (plotted points in Figure 12). Hence, we can conclude that the patients not having died or relapsed after 1000 days of observation were at negligible risk of death or relapse. Further, only two exits from stage 6 occurred after 600 days, so that the risk of relapse or death 600 days after chronic GVHD onset was low.

The regression coefficient $\hat{B}_{16}(t)$ progressed in a negative direction, indicating that patients with longer waiting times to chronic GVHD onset experienced a reduced hazard of death or relapse (i.e. longer stage 6 waiting times) after chronic GVHD onset. This effect was apparently significant, judging by the pointwise 95% confidence limits which excluded

the null value 0 beyond 200 days after chronic GVHD onset. A normal test of the integrated regression coefficient $B_{16}(t)$ at the end of observation (day 2102, the last censoring time) was strongly significant ($z = 3.59, p = .0003$). Patient age at transplantation had no apparent effect on the risk of death or relapse secondary to chronic GVHD. The integrated coefficient function $\hat{B}_{26}(t)$ was non-monotonic and remained close to zero, with large

Figure 12. Estimates of integrated regression coefficients from model (4.11) of bone marrow transplant data. The IPCW estimate is provided by a solid line, the naive estimate by a dotted line, and approximate pointwise 95% CI for the IPCW estimate by dashed lines. Plotted points are censored waiting times.



confidence intervals containing 0 over all observed waiting times. The normal test of $\hat{B}_{26}(t)$ was non-significant ($z = 0.63, p = 0.53$). The low risk and high risk AML groups were at greater risk of death or relapse secondary to chronic GVHD, as shown by the positive integrated regression coefficients $\hat{B}_{36}(t)$ and $\hat{B}_{46}(t)$. However, these differences were small

relative to the variability of the hazard estimates and were non-significant based on a normal test (low risk AML: $z = 1.31, p = .19$; high risk AML: $z = 0.92, p = .36$). Additionally, a comparison of all three groups was non-significant ($\chi_1^2 = 2.57, p = .11$).

The IPCW regression coefficient estimates did not substantially differ from the naive estimates (dotted lines, Figure 12). The maximum differences between the IPCW and naive estimates of the five integrated regression coefficients (intercept included) were 0.12, 0.0003, 0.0019, 0.06, and 0.02, respectively. These maximal differences occurred at or close to the largest stage 6 waiting times and represented less than 10% of the IPCW coefficient estimates. These small differences in estimates suggested that any dependencies among censoring and transition times had a minimal effect on our marginal analysis of stage 6 waiting times.

From a clinical perspective, our results demonstrated that neither patient age nor disease group (ALL, low risk AML, high risk AML) were significant determinants of the risk of death or relapse following the onset of chronic GVHD. Previous analyses [25] of these data demonstrated that low risk AML patients exhibited significantly lower hazard of death or relapse than ALL and high risk AML patients, both marginally and after adjusting for patient age (which was non-significantly associated with the death/relapse hazard). However, the results of this analysis have little bearing on our present analysis, as it considered death/relapse hazards after transplantation, (i.e.) the hazard of death/relapse from stage 1. We have additionally demonstrated that patients who more rapidly developed chronic GVHD exhibited significantly greater hazard of death or relapse after chronic GVHD onset. It would be reasonable to speculate, then, that patients with rapid onset chronic GVHD should be closely monitored for the risk of death or relapse. We note that there was some covariation among our selected predictors. Among the 59 patients that developed chronic GVHD, patient ages significantly differed (Kruskal-Wallis test, $p = .005$), with the ALL group exhibiting lower ages (mean age = 23) than the low and high risk AML groups (mean ages 30 and 31, respectively). Further, the time to onset of chronic GVHD was significantly different among groups (Kruskal-Wallis test, $p = .02$) – ALL patients most rapidly developed chronic GVHD (median wait = 120 days), followed by high risk AML patients (150 days) and low risk AML patients (180 days). Patient age and time to chronic GVHD onset were marginally significantly related (Spearman correlation = 0.23, $p = .07$). However, our empirical inspection of the association between stage 6 waiting times (Figure 11) and our model covariates seemed to illustrate no marginal

association between patient age, disease group, and stage 6 waiting times, and we so suspect no confounding relationships among our covariates.

CHAPTER V

CONCLUSION

In this research, we have applied the principle of inverse probability of censoring weighting (IPCW) to develop K -sample test statistics and nonparametric regression coefficient estimators for waiting times from multi-stage models. In Chapter III, we introduced K -sample test statistics for waiting times in a multi-stage model that were analogous to traditional test statistics for survival data, such as the log-rank test. The simulation studies we conducted demonstrated that our test statistics, via the IPCW approach, were robust against violations of the typically-required assumptions of independent censoring and the semi-Markov property. Further, the simulations demonstrated that test statistics for survival data were generally inappropriate for waiting times from a multi-stage model, exhibiting inflated size and non-monotonic relationships with the effect size implied by the alternative hypothesis. The simulations also illustrated that our test statistic exhibited greater power than a generalized IPCW Mann-Whitney statistic for waiting times under proportional hazards alternatives. This corresponded with the well-known result for survival data that log-rank tests are optimal for proportional hazards alternatives.

In Chapter IV, we introduced a nonparametric regression model for waiting times corresponding to Aalen's linear model for survival data. Our simulation studies in Chapter IV demonstrated that the IPW coefficient estimator was approximately unbiased in several situations, particularly under violations of the semi-Markov property. Further, as expected, the unadjusted, naive coefficient estimators exhibited substantial bias when waiting times from successive stages in the multi-stage model were correlated. These results were contemporary with the results of the simulation study of the hypothesis test statistic – that IPCW statistics perform adequately under violations of the typical assumptions made of multi-stage models and the statistics adapted from survival data can be significantly biased.

A significant advantage of the IPCW approach is its relative simplicity. As frequently noted above, the IPCW principle itself is quite simple – replacing the

uncensored data counting processes used for traditional estimators with their IPCW equivalents. Hence, armed with reasonable estimates of the probability of censoring, IPCW estimators and test statistics are fairly simple to construct. However, calculating estimates of the probability of censoring to use as weights can be fairly complicated and requires some attention. A detailed analysis of censoring patterns should be conducted before generating the weights for IPCW statistics, particularly in the context of a multi-stage model. Such an analysis would ensure that as accurate an estimate of censoring probabilities is generated. In some capacity, the complexity of such a preliminary analysis of censoring patterns is related to the complexity of the multi-stage network, particularly if stage-dependent censoring is suspected. Nevertheless, the IPCW statistics we have introduced were of closed form and calculable from the available data as were estimates of variance, regardless of the observed censoring patterns.

Another advantage of the IPCW approach is its generality. The IPCW principle of replacing uncensored data counting processes with IPCW analogues applies directly to waiting times from multi-stage data just as with survival data, for which it was initially developed. The IPCW method does not make unnecessary assumptions about censoring patterns nor the independence of transition times from a multi-stage model. Further, the proof of the asymptotic equivalence of IPCW counting processes and uncensored data counting processes was independent of the method used to estimate the probability of censoring, (i.e.) the proofs were based on any estimate of the probability of censoring rather than a particular estimates. Therefore, any asymptotically consistent estimator of the probability of censoring can be used. Our use of Aalen's linear model for the censoring hazard was predicated by the unique flexibility it possesses via its coefficient functions. We do recall that the selection of estimate for the probability of censoring does play a role in the asymptotic distribution of IPCW statistics, but this dependence largely impacted expressions for the asymptotic variance and estimators thereof.

Aalen's linear model has received only minimal attention, and to our knowledge the results of Chapter IV are the first attempt at extensions under dependent censoring or for multi-stage models. As noted in the introduction, there are several methods for hypothesis testing of survival times and waiting times from a multi-stage model under dependent censoring. We note that the test statistics of Lin, Robins, and Wei [31], Lin and Ying [33], and Fan and Datta [15] were developed explicitly for two populations, whereas our proposed test statistics apply to $K > 2$ populations. The regression estimators by several

authors detailed in the introduction can provide tests for $K > 2$ populations as well. The IPCW U -statistic proposed by Datta, et al [10] provide a very general result with broad applicability. In the same way many common statistics from classical statistics can be expressed as traditional U -statistics, we expect that many IPCW statistics and estimators have representation as IPCW U -statistics.

As noted in Chapter II, the IPCW principle has been previously applied to develop estimators of the cumulative hazard and survival functions for waiting times in a multi-stage model [44]. In this research, we have developed K -sample log-rank-type test statistics and nonparametric regression estimators. Each of these statistics are analogous to their survival data counterparts – the Nelson-Aalen estimator for cumulative hazards, the Kaplan-Meier estimator for survival functions, the log-rank test, and Aalen’s linear model. As noted earlier, perhaps the most popular regression model for censored survival data is the Cox proportional hazards model [9]. The Cox model specifies a multiplicative hazard function, in which the hazard of an event is modeled as the product of a “baseline” hazard function and an exponentiated linear combination of model covariates. While IPCW estimators akin to those from the Cox model have been developed for censored survival data, extensions to waiting times from multi-stage models have not been pursued. This provides an avenue for further research, and we briefly motivate it below.

In our notation, the Cox model adapted to waiting times from a multi-stage model specifies the multiplicative hazard function

$$\lambda_i(t, \mathbf{X}_i(t)) = \lambda_{0j}(t) \exp\left(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij}^* + t)\right), \quad (5.1)$$

where $\lambda_{0j}(t)$ denotes a baseline hazard for stage j exits and $\boldsymbol{\beta}$ is the vector of regression coefficients. (A more general formulation, referred to as a “relative risk” model, replaces the exponential function with a general risk function $r(\boldsymbol{\beta}, \mathbf{X}_i)$.) Again, we note that the covariate vector is evaluated at $T_{ij}^* + t$ since the covariate processes evolve in calendar time but our interest is in modeling what occurs after stage j entry. The baseline hazard $\lambda_{0j}(t)$ is left completely unspecified. Omitting the details of the derivation (see Andersen, et al [6] for a detailed treatment), the coefficient vector $\boldsymbol{\beta}$ is estimated as a solution to the score equation $\mathbf{U}_\tau^*(\boldsymbol{\beta}) = 0$, where

$$\mathbf{U}_\tau^*(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^\tau \mathbf{X}_i(T_{ij}^* + t) dN_{ij}^*(t) - \int_0^\tau \mathbf{E}(\boldsymbol{\beta}, t) dN_j^*(t), \quad (5.2)$$

where the covariate vectors $\mathbf{X}_i(t)$ and stage j uncensored data counting process $N_{ij}^*(t)$ and $N_j^*(t)$ are as defined in the text, τ represents a sufficiently large time point to accumulate all exits from stage j , and the vector $\mathbf{E}^*(\boldsymbol{\beta}, t)$ is

$$\mathbf{E}^*(\boldsymbol{\beta}, t) = \frac{\mathbf{S}^{(1)*}(\boldsymbol{\beta}, t)}{S^{(0)*}(\boldsymbol{\beta}, t)}, \quad (5.3)$$

where

$$\begin{aligned} S^{(0)*}(\boldsymbol{\beta}, t) &= \sum_{i=1}^n \exp\left(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij}^* + t)\right) Y_{ij}^*(t) \\ \mathbf{S}^{(1)*}(\boldsymbol{\beta}, t) &= \sum_{i=1}^n \mathbf{X}_i(T_{ij}^* + t) \exp\left(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij}^* + t)\right) Y_{ij}^*(t), \end{aligned} \quad (5.4)$$

and the Y_{ij}^* are the uncensored data stage j at risk processes. The covariance matrix of the estimator $\boldsymbol{\beta}^*$ arrived at via the score equation (5.2) can be estimated by $\mathbf{I}_\tau^*(\boldsymbol{\beta})$, where

$$\mathbf{I}_\tau^*(\boldsymbol{\beta}) = \int_0^\tau \mathbf{V}^*(\boldsymbol{\beta}, t) dN_j^*(t), \quad (5.5)$$

where $\mathbf{V}^*(\boldsymbol{\beta}, t) = \mathbf{S}^{(2)*}(\boldsymbol{\beta}, t)/S^{(0)*}(\boldsymbol{\beta}, t) - \mathbf{E}^*(\boldsymbol{\beta}, t)\mathbf{E}^*(\boldsymbol{\beta}, t)^T$, and

$$\mathbf{S}^{(2)*}(\boldsymbol{\beta}, t) = \sum_{i=1}^n \mathbf{X}_i(T_{ij}^* + t)\mathbf{X}_i(T_{ij}^* + t)^T \exp\left(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij}^* + t)\right) Y_{ij}^*(t). \quad (5.6)$$

Finally, the baseline cumulative hazard $\Lambda_{0j}(t) = \int_0^t \lambda_{0j}(s) ds$ can be estimated by

$$\Lambda_{0j}^*(t) = \int_0^t \frac{J_j^*(s)}{S^{(0)*}(\boldsymbol{\beta}, s)} dN_j^*(s), \quad (5.7)$$

where $J_j^*(s) = I[Y_j^*(s) > 0]$.

Under suitable regularity conditions on the functions $S^{(0)*}$, $\mathbf{S}^{(1)*}$, $\mathbf{S}^{(2)*}$, \mathbf{E}^* , and \mathbf{V}^* , it can be shown that a solution $\boldsymbol{\beta}^*$ to $\mathbf{U}_\tau^*(\boldsymbol{\beta}) = 0$ exists and is unique with probability approaching 1 and $\boldsymbol{\beta}^*$ converges in probability to the “true” value of the covariate vector (see Condition VII.2.1 and Theorem VII.2.1 of Andersen, et al [6]). Under an additional Lindeberg assumption on the covariate processes $\mathbf{X}_i(t)$, the covariate estimator $\boldsymbol{\beta}^*$ can be shown to be asymptotically normal with covariance matrix consistently estimated by $n^{-1}\mathbf{I}_\tau^*(\boldsymbol{\beta}^*)$ (see Condition VII.2.2 and Theorem VII.2.2 of Andersen, et al [6]). A similar weak convergence result holds for the estimator (5.7) of the cumulative baseline hazard function.

As expected, the applicability of the Cox model for right censored survival data hinges in some part on the independence of censoring times from survival times. Further, a waiting time Cox model would require independent censoring and the semi-Markov property. We therefore propose IPCW estimators of the model coefficients and baseline cumulative hazard function for waiting times in a multi-stage model. As before, the definition of these estimators is simply the replacement of uncensored data counting processes in the uncensored data estimators with their IPCW equivalents. In detail, define the following quantities:

$$\begin{aligned}
\widehat{S}^{(0)}(\boldsymbol{\beta}, t) &= \sum_{i=1}^n \exp(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij} + t)) \widehat{Y}_{ij}(t) \\
\widehat{S}^{(1)}(\boldsymbol{\beta}, t) &= \sum_{i=1}^n \mathbf{X}_i(T_{ij} + t) \exp(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij} + t)) \widehat{Y}_{ij}(t) \\
\widehat{S}^{(2)}(\boldsymbol{\beta}, t) &= \sum_{i=1}^n \mathbf{X}_i(T_{ij} + t) \mathbf{X}_i(T_{ij} + t)^T \exp(\boldsymbol{\beta}^T \mathbf{X}_i(T_{ij} + t)) \widehat{Y}_{ij}(t) \\
\widehat{\mathbf{E}}(\boldsymbol{\beta}, t) &= \frac{\widehat{S}^{(1)}(\boldsymbol{\beta}, t)}{\widehat{S}^{(0)}(\boldsymbol{\beta}, t)} \\
\widehat{\mathbf{U}}_\tau(\boldsymbol{\beta}) &= \sum_{i=1}^n \int_0^\tau \mathbf{X}_i(T_{ij} + t) d\widehat{N}_{ij}(t) - \int_0^\tau \widehat{\mathbf{E}}(\boldsymbol{\beta}, t) d\widehat{N}_j(t), \tag{5.8}
\end{aligned}$$

where $\widehat{N}_{ij}, \widehat{N}_j, \widehat{Y}_{ij}, \widehat{Y}_j$ are the stage j exit and at risk IPCW counting process defined before. We then suggest IPCW estimators $\widehat{\boldsymbol{\beta}}$ as the solutions to the score equation $\widehat{\mathbf{U}}_\tau(\boldsymbol{\beta}) = 0$. We additionally suggest an IPCW estimator for the baseline hazard function

$$\widehat{\Lambda}_{0j}(t) = \int_0^t \frac{\widehat{J}_j(s)}{\widehat{S}^{(0)}(\boldsymbol{\beta}, s)} d\widehat{N}_j(s), \tag{5.9}$$

where $\widehat{J}_j(s) = I[\widehat{Y}_j(s) > 0]$.

The consistency and asymptotic normality proofs for the estimator $\boldsymbol{\beta}^*$ for uncensored data partially rest upon a martingale representation for the score function. In particular, the k^{th} component of the score vector $\mathbf{U}_t(\boldsymbol{\beta})$ has representation

$$U_t^{k*}(\boldsymbol{\beta}) = \sum_{i=1}^n \int_0^t (X_{ik}(T_{ij}^* + s) - E_k^*(\boldsymbol{\beta}_0, s)) dM_i^*(s), \tag{5.10}$$

where the X_{ik} and E_k^* are the components of \mathbf{X}_i and \mathbf{E}^* , respectively, and $k = 1, \dots, p$. A related martingale representation for the baseline hazard estimator is also the foundation for the proof of its asymptotic normality. We expect that a martingale representation for

the IPCW score function $\widehat{U}_\tau(\beta)$ can be achieved and asymptotic results derived therefrom. The martingale representations we derived for IPCW statistics in this dissertation decomposed into the sum of two orthogonal martingales – (1) a stochastic integral of a predictable process with respect to the martingale associated with the uncensored data counting process for stage j exits, and (2) a stochastic integral of a predictable process with respect to the martingale associated with the counting process for censoring events. Further, the first martingale of these decompositions corresponded directly to the martingale associated with a given uncensored data estimator/test statistic. Based on these results, we expect that such a decomposition could be achieved for the IPCW score process $\widehat{U}_\tau(\beta)$, and that the first summand of said decomposition would be precisely the score process $U_\tau^*(\beta)$. Further, we expect that the asymptotic variance of IPCW Cox model-type estimators would be the sum of the uncensored data variance expression $I_\tau^*(\beta)$ and the predictable covariation process of the second martingale of the decomposition.

For asymptotic considerations, the conditions required for the consistency of $\widehat{\beta}$ would include those conditions required for β^* , but would likely require additionally assumptions guaranteeing the asymptotic negligibility of the second martingale from the decomposition. Similarly, the conditions required for asymptotic normality of $\widehat{\beta}$ would include those for β^* but would require additional assumptions on the second martingale. Practically speaking, these weak convergence assumptions will likely be a function of the method of estimation for the inverse probability of censoring. For our results, they were regularity conditions on the vector of covariates used to estimate the censoring hazard via Aalen’s linear model. These concepts, of course, need more formalization and investigation, but do provide a fruitful direction for additional research into inference for multi-stage models under dependent censoring and violations of the semi-Markov property.

REFERENCES

- [1] Aalen OO. Nonparametric inference for a family of counting processes. *Ann. Statist.* 1978; 6:701-726.
- [2] Aalen OO. A model for nonparametric regression analysis of counting processes. In *Lecture Notes on Mathematical Statistics and Probability*, 2, Klonecki W, Kozek A, Rosiski J (eds). Springer-Verlag: NY, 1980; 1-25.
- [3] Aalen OO. A linear regression model for the analysis of lifetimes. *Statist. Med* 1989; 8:907-925.
- [4] Aalen OO. Further results on the nonparametric linear regression model in survival analysis. *Statist. Med.* 1993; 12, 1569-1588.
- [5] Aalen OO, Johansen S. An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scand. J. Statist.* 1978; 5, 141-150.
- [6] Andersen PK, Borgan Ø, Gill RD, Keiding N. *Statistical Models Based on Counting Processes*. Springer-Verlag, NY, 1993.
- [7] Cain LE, Cole SR. Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death. *Statist. Med* 2009; 28(12):1725-38.
- [8] Copelan EA, Biggs JC, Thompson JM, Crilley P, Szer J, Klein JP, Kapoor N, Avalos BR, Cunningham I, Atkinson K. Treatment for acute myelocytic leukemia with allogeneic bone marrow transplantation following preparation with BuCy2. *Blood*, Aug 1991; 78: 838-843.
- [9] Cox DR. Regression models and life tables (with discussion). *J. Roy. Statist. Soc. B* 1972; 34: 187-220.
- [10] Datta S, Bandyopadhyay D, Satten GA. Inverse probability of censoring weighted

- U*-statistics for right-censored data with an application to testing hypotheses. *Scand J Statist* 2010; doi: 10.1111/j.1467-9469.2010.00697.x.
- [11] Datta S, Satten GA. Nonparametric estimation of stage occupation probabilities and transition hazards under dependent censoring. *University of Georgia Department of Statistics Technical Report*. Stat 2000-4.
- [12] Datta S, Satten GA. Validity of the Aalen-Johansen estimators of stage occupation probabilities and Nelson-Aalen estimators of integrated transition hazards for non-Markov models. *Stat Prob Lett* 2001; 55: 403-411.
- [13] Datta S, Satten GA. Estimation of integrated transition hazards and stage occupation probabilities for non-Markov systems under dependent censoring. *Biometrics* 2002;58(4):792-802.
- [14] Efron B, Tibshirani. *Introduction to the Bootstrap*. London: Chapman and Hall, 1993.
- [15] Fan J, Datta S. On Mann-Whitney tests for comparing sojourn time distributions when the transition times are right censored. *Submitted* 2010.
- [16] Harrington DP, Fleming TR. A class of rank test procedures for censored survival data. *Biometrika* 1982; 69, 553-566.
- [17] Grønnesby JK. Testing covariate effects in Aalen's linear model. *Scand J Statist*. 1997; 24(1): 125-35.
- [18] Gunnes N, Borgan O, Aalen OO. Estimating stage occupation probabilities in non-Markov models. *Lifetime Data Anal*. 2007; 13(2):211-40.
- [19] Horvitz DG and Thompson DJ. A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc*. 1952; 47, 663-685.
- [20] Hosmer DW and Royston P. Using Aalens linear hazards model to investigate time-varying effects in the proportional hazards regression model. *The Stata Journal* 2002; 2 (4), 331-350.
- [21] Howe CJ, Cole SR, Chmiel JS, Muñoz A. Limitation of Inverse Probability-of-Censoring Weights in Estimating Survival in the Presence of Strong Selection Bias. *Am J Epidem* 2011; DOI: 10.1093/aje/kwq385.

- [22] Huang Y. Censored regression with the multistate accelerated sojourn times model. *Journal of the Royal Statistical Society Series B* 2002; 64, 17-29.
- [23] Ichida JM, Wassel JT, Keller MD, Ayers LW. Evaluation of protocol change in burn-care management using the Cox proportional hazards model with time-dependent covariates. *Statist. Med.* 1993; 12, 301-310.
- [24] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* 1958; 53, 457-481.
- [25] Klein JP, Moeschberger ML. *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer Verlag, 1997.
- [26] Klein JP, Moeschberger ML. Data sets from *Survival Analysis: Techniques for Censored and Truncated Data*, New York: Springer Verlag, 1997.
<http://www.mcw.edu/biostatistics/Faculty/Faculty/JohnPKleinPhD/SurvivalAnalysisBook/DataSetsBothEditions.htm>.
 First retrieved October 1, 2010.
- [27] Klein JP, Moeschberger ML. Modified by Yan J. Data sets from Klein and Moeschberger (1997), *Survival Analysis*. R package version 0.1-3. 2003.
- [28] Koul H, Susarla V, Van Ryzin J. Regression analysis with randomly right censored data. *Annals of Statistics* 1981; 9, 1276-1288.
- [29] Lakhal-Chaieb L, Cook RJ, Lin X. Inverse Probability of Censoring Weighted Estimates of Kendall's τ for Gap Time Analyses. *Biometrics* 2010; 66(4), 1145-52.
- [30] Lee E, Weissfeld LA. Assessment of covariate effects in Aalen's additive hazard model. *Statist. Med.* 1998; 17(9), 983-98.
- [31] Lin DY, Robins JM, Wei LJ. Comparing two failure time distributions in the presence of dependent censoring. *Biometrika* 1996; 83(2), 381-393.
- [32] Lin DY, Sun W, Ying Z. Nonparametric estimation of the gap time distributions for serial events with censored data. *Biometrika* 1999; 86, 59-70.
- [33] Lin DY, Ying Z. Nonparametric tests for the gap time distributions of serial events based on censored data. *Biometrics* 2001; 57, 369-375.

- [34] Mantel N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports* 1996; 50 (3), 163-70.
- [35] Matsuyama Y, Yamaguchi T. Estimation of the marginal survival time in the presence of dependent competing risks using inverse probability of censoring weighted (IPCW) methods. *Pharm Stat* 2008; 7(3):202-14.
- [36] Nelson W. Theory and applications of hazard plotting for censored failure data. *Technometrics* 1972; 14:945-966.
- [37] Pepe MS, Fleming TR. Weighted Kaplan-Meier statistics: A class of distance tests for censored survival data. *Biometrics* 1989; 45, 497-507.
- [38] Robins J, Finkelstein D. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* 2000; 56, 779-788.
- [39] Robins JM and Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology – Methodological Issues*, N. Jewell, K. Dietz, and V. Farewell (eds), 297-331. Boston: Birkhauser, 1993.
- [40] Robins J, Rotnitzky A. Semiparametric regression estimation in the presence of dependent censoring. *Biometrika* 1995; 82, 805-820.
- [41] Satten GA, Datta S. The Kaplan-Meier estimator as an inverse-probability-of-censoring weighted average. *The American Statistician* 2000; 55, 207-210.
- [42] Satten GA, Datta S, Robins J. An estimator for the survival function when data are subject to dependent censoring. *University of Georgia Department of Statistics Technical Report*. Stat 2001-7.
- [43] Satten GA, Datta S, Robins J. Estimating the marginal survival function in the presence of time dependent covariates. *Stat Prob Lett* 2001; 54: 397-403.
- [44] Satten GA, Datta S. Marginal estimation for multi-stage models: waiting time distributions and competing risks analyses. *Statist. Med.* 2002; 21: 3-19.
- [45] Schaubel DE, Cai J. Regression methods for gap time hazard functions of sequentially ordered multivariate failure time data. *Biometrika* 2003; 91, 291-303.

- [46] Strawderman RL. The accelerated gap times model. *Biometrika* 2005; 92, 647-666.
- [47] Wang W, Wells MT. Nonparametric estimation of successive duration times under dependent censoring. *Biometrika* 1998; 85, 561-572.

APPENDIX

Proofs of Lemmas and Theorems

Proof of Lemma II.4.

We first note that $\widehat{\Lambda}_j(t) - \Lambda_j(t)$ can be decomposed as follows:

$$\begin{aligned} \widehat{\Lambda}_j(t) - \Lambda_j(t) &= \int_0^t \frac{1}{\widehat{Y}_j(s)} d(\widehat{N}_j(s) - N_j^*(s)) + \int_0^t \left(\frac{1}{\widehat{Y}_j(s)} - \frac{1}{Y_j^*(s)} \right) dN_j^*(s) + \\ &\quad \int_0^t \frac{dN_j^*(s)}{Y_j^*(s)} - \Lambda_j(t). \end{aligned} \quad (\text{A.1})$$

Denote the three summands of the right hand side of (A.1) sum as I, II, and III. By standard martingale methods, we then have that

$$\text{III} = \int_0^t \frac{J_j^*(s)}{Y_j^*(s)} dM_j^*(s) + o_p(n^{-1/2}), \quad (\text{A.2})$$

where where $J_j^*(s) = I[Y_j^*(s) > 0]$. We note that $\frac{1}{n}\widehat{Y}_j(t) = S_j(t) + o_p(n^{-1/2})$, by the following chain of argument: (1) $\widehat{Y}_j(t) \xrightarrow{p} \bar{Y}_j(t)$ by the consistency of $\widehat{\mathbf{B}}(t)$ under Aalen's linear model, (2) $\bar{Y}_j(t) = Y_j^*(t) + o_p(n^{-1/2})$ by Theorem II.1, (3) $\frac{1}{n}Y_j^*(t) \xrightarrow{p} S_j(t)$ by the law of large numbers. Hence, we have that the first term satisfies

$$\begin{aligned} \text{I} &= \int_0^t \frac{1}{\widehat{Y}_j(s)} d(\widehat{N}_j(s) - N_j^*(s)) \\ &= \int_0^t \frac{1}{nS_j(s)} d(\widehat{N}_j(s) - N_j^*(s)) + o_p(n^{-1/2}) \\ &= \sum_{i=1}^n \frac{1}{nS_j(U_{ij}^* - T_{ij}^*)} \left(\frac{I[U_{ij}^* - T_{ij}^* \leq t] \delta_{ij}}{\widehat{K}_i(U_{ij}^-)} - I[U_{ij}^* - T_{ij}^* \leq t, X_{ij} = 1] \right) + o_p(n^{-1/2}) \\ &= \sum_{i=1}^n \frac{I[U_{ij}^* - T_{ij}^* \leq t, X_{ij} = 1]}{nS_j(U_{ij}^* - T_{ij}^*)} \left(\frac{I[C_i \geq U_{ij}^*]}{\widehat{K}_i(U_{ij}^-)} - 1 \right) + o_p(n^{-1/2}). \end{aligned}$$

Using Lemma II.2 (and the note directly following it), the consistency of \widehat{K}_i , and the law of

large numbers, we get

$$\begin{aligned}
\text{I} &= - \sum_{i=1}^n \frac{I[U_{ij}^* - T_{ij}^* \leq t] \Phi_j}{n S_j(U_{ij}^* - T_{ij}^*)} \int_0^{U_{ij}^* -} \frac{1}{K_i(s)} d\widehat{M}_i^c(s) + o_p(n^{-1/2}) \\
&= - \sum_{i=1}^n \frac{I[U_{ij}^* - T_{ij}^* \leq t] \Phi_j}{n S_j(U_{ij}^* - T_{ij}^*)} \int_0^\infty \frac{I[s < U_{ij}^*]}{K_i(s)} d\widehat{M}_i^c(s) + o_p(n^{-1/2}). \tag{A.3}
\end{aligned}$$

We next note that $\frac{1}{n} \widehat{N}_j(t) = 1 - S_j(t) + o_p(n^{-1/2})$ by the same chain of argument asserting that $\frac{1}{n} \widehat{Y}_j(t) = S_j(t) + o_p(n^{-1/2})$. We now observe that

$$\begin{aligned}
\text{II} &= \int_0^t \frac{1}{n S_j^2(s)} (\widehat{Y}_j(s) - Y^*(s)) dS_j(s) + o_p(n^{-1/2}) \\
&= \sum_{i=1}^n \int_0^t \frac{I[X_{ij} = 1]}{n S_j^2(s)} \left(\frac{I[U_{ij}^* - T_{ij}^* \geq s, C_i \geq T_{ij}^*]}{\widehat{K}_i(T_{ij}^* + s-)} - 1 \right) dS_j(s) + o_p(n^{-1/2}).
\end{aligned}$$

By Lemma II.2, the consistency of \widehat{K}_i , and the law of large numbers we have

$$\text{II} = - \sum_{i=1}^n \int_0^t \frac{I[U_{ij}^* - T_{ij}^* \geq s] \Phi_j}{n S_j^2(s)} \int_0^{T_{ij}^* + s-} \frac{1}{K_i(u)} d\widehat{M}_i^c(u) dS_j(s) + o_p(n^{-1/2}).$$

Interchanging the order of integration on the right hand side produces

$$\begin{aligned}
\text{II} &= \sum_{i=1}^n \int_0^{T_{ij}^* + t-} \left\{ \int_{0 \vee (s - T_{ij}^*)}^t \frac{I[U_{ij}^* - T_{ij}^* \geq u] \Phi_j}{n S_j^2(u)} dS_j(u) \right\} \frac{1}{K_i(s)} d\widehat{M}_i^c(s) + o_p(n^{-1/2}) \\
&= \sum_{i=1}^n \int_0^\infty g_{ij}(s, t) d\widehat{M}_i^c(s) + o_p(n^{-1/2}),
\end{aligned}$$

where

$$g_{ij}(s, t) = \frac{\Phi_j}{n K_i(s)} \left(\frac{I[s - t < T_{ij}^*]}{S_j(t \wedge (U_{ij}^* - T_{ij}^*))} - \frac{I[s - t < T_{ij}^*]}{S_j((0 \vee (s - T_{ij}^*)) \wedge (U_{ij}^* - T_{ij}^*))} \right). \tag{A.4}$$

Combining the expressions (A.3) and (A.4) for terms I and II of (A.1) produces

$$\text{I} + \text{II} = \sum_{i=1}^n \int_0^\infty \zeta_{ij}(s, t) d\widehat{M}_i^c(s) + o_p(n^{-1/2}),$$

where $\zeta_{ij}(s, t)$ is as previously defined in (2.4). By the definitions of $M_i^c(s)$ and $\widehat{M}_i^c(s)$, we

can write $d\widehat{M}_i^c(s) = dM_i^c(s) - I[T_i \geq s] d[\widehat{\Lambda}_i^c - \Lambda_i^c](s)$, so that

$$\text{I} + \text{II} = \sum_{i=1}^n \left\{ \int_0^\infty \zeta_{ij}(s, t) dM_i^c(s) - \int_0^\infty \zeta_{ij}(s, t) I[T_i \geq s] d[\widehat{\Lambda}_i^c - \Lambda_i^c](s) \right\} + o_p(n^{-1/2}). \quad (\text{A.5})$$

From Aalen's linear model and our choice of generalized inverse for the matrix $\mathbf{A}(t)$, we have that for $T_i \geq s$,

$$\widehat{\Lambda}_i^c(s) - \Lambda_i^c(s) = \int_0^s \mathbf{X}_i^T(u) \mathbf{A}^{-1}(u) \widetilde{\mathbf{X}}^T(u) d\mathbf{M}^c(u) + o_p(n^{-1/2}),$$

where $\mathbf{X}_i(u)$, $\widetilde{\mathbf{X}}_i(u)$, $\widetilde{\mathbf{X}}(u)$, $\mathbf{A}(u)$, and $\mathbf{M}^c(u)$ are as defined in the text. We then have that (A.5) becomes

$$\text{I} + \text{II} = \int_0^\infty \zeta_j^T(s, t) [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s)] d\mathbf{M}^c(s) + o_p(n^{-1/2}).$$

Finally, adding in expression (A.2) for term III, we arrive that

$$\begin{aligned} \widehat{\Lambda}_j(t) - \Lambda_j(t) &= \int_0^t \frac{J_j^*(s)}{Y_j^*(s)} dM_j^*(s) + \\ &\quad \int_0^\infty \zeta_j^T(s, t) [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s)] d\mathbf{M}^c(s) + o_p(n^{-1/2}), \end{aligned}$$

the desired martingale representation, thus completing the proof of Lemma II.4. ■

Proof of Theorem III.1.

We first recall the martingale representation and predictable covariation process of $\widehat{\mathbf{Z}}_j(t)$,

$$\begin{aligned} \widehat{Z}_{hj}(t) &= \sum_{k=1}^K \int_0^t \widehat{W}_j(s) \left(\Delta_{hk} - \frac{Y_{kj}^*(s)}{Y_j^*(s)} \right) dM_{kj}^*(s) + \\ &\quad \int_0^\infty \boldsymbol{\eta}_j^K(s, t)^T [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s)] d\mathbf{M}^c(s) + o_p(n^{-1/2}), \end{aligned} \quad (\text{A.6})$$

$$\begin{aligned} \langle \widehat{Z}_{hj}(t), \widehat{Z}_{kj}(t) \rangle &= \int_0^t \widehat{W}_j^2(s) Y_{hj}^*(s) \left(\Delta_{hk} - \frac{Y_{kj}^*(s)}{Y_j^*(s)} \right) \lambda_j(s) ds + \\ &\quad \int_0^\infty \boldsymbol{\eta}_j^K(s, t)^T [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s)] \times \mathbf{I}_\lambda(s) \times \end{aligned}$$

$$\left[\mathbf{I} - \widetilde{\mathbf{X}}(s)\mathbf{A}^{-1}(s)\widetilde{\mathbf{X}}^T(s) \right] \boldsymbol{\eta}_j^K(s, t) ds + o_p(n^{-1/2}), \quad (\text{A.7})$$

Given the orthogonality [43] of the two terms of the martingale representation of the h^{th} component of $\widehat{\mathbf{Z}}_j(t)$ in (A.6), we prove Theorem III.1 by demonstrating the weak convergence of each term to a normal limit. The first term,

$$\sqrt{n} \sum_{k=1}^K \int_0^t \widehat{W}_j(s) \left(\Delta_{hk} - \frac{Y_{hj}^*(s)}{Y_{kj}^*(s)} \right) dM_{kj}^*(s), \quad (\text{A.8})$$

directly corresponds to the martingale representation of K -sample test statistics for counting processes as detailed in [6] (cf. Equation 5.2.3, p. 346). The conditions required for weak convergence defined therein (cf. Theorem V.2.1, p. 360) are the existence of deterministic functions $y_{1j}^*, \dots, y_{Kj}^*, y_{.j}^* \psi_j$ such that for all h, k, t

- $\psi_j^2 y_{.j}^* \lambda_j$ is integrable,
- $n \int_0^t \widehat{W}_j^2(s) Y_{hj}^*(s) Y_{kj}^*(s) (Y_{.j}^*(s))^{-1} \lambda_j(s) ds \xrightarrow{p} \int_0^t \psi_j^2(s) y_{hj}^*(s) y_{kj}^*(s) (y_{.j}^*(s))^{-1} \lambda_j(s) ds$,
(convergence of the predictable covariation process to a continuous, deterministic limit)
- $n \int_0^t \widehat{W}_j^2(s) I \left[\left| \sqrt{n} \widehat{W}_j(s) \right| > \epsilon \right] Y_{.j}^*(s) \lambda_j(s) ds \xrightarrow{p} 0$. (Lindeberg condition)

Under these assumptions, the conditions of the Rebelledo martingale central limit theorem are satisfied, and the K -vector of terms (A.8) converges in distribution to a normal limit with covariance defined by the first integral of (A.7) [6].

We now consider the second term of (A.6) for the h^{th} component of $\sqrt{n} \widehat{\mathbf{Z}}_j(t)$,

$$\sqrt{n} \int_0^\infty \boldsymbol{\eta}_j^h(s, t)^T \left(\mathbf{I} - \widetilde{\mathbf{X}}(s)\mathbf{A}^{-1}(s)\widetilde{\mathbf{X}}^T(s) \right) d\mathbf{M}^c(s). \quad (\text{A.9})$$

Both $\zeta_j^K(s, t)$ and $\zeta_j(s, t)$ in the definition of $\boldsymbol{\eta}_j^K(s, t)$ are deterministic. Further, up to $o_p(n^{-1/2})$ terms, $\widehat{Y}_{hj}(t) = n_h S_{hj}(t)$; hence, we write

$$\boldsymbol{\eta}_j^K(s, t) = n_h \int_0^t \widehat{W}_j(u) S_{hj}(u) \left(\zeta_j^K(s, du) - \zeta_j(s, du) \right) + o_p(n^{-1/2}).$$

We now make two assumptions regarding the asymptotic behavior of $\boldsymbol{\eta}_j^K(s, t)$, in particular imposing further conditions on the function ψ_j defined above such that for all h ,

- $\psi_j(u) S_{hj}(u) \left(\zeta_j^K(s, du) - \zeta_j(s, du) \right)$ is integrable with respect to u ,

- $n_h \int_0^t \widehat{W}_j(u) S_{hj}(u) \left(\zeta_j^K(s, du) - \zeta_j(s, du) \right) \xrightarrow{p} \int_0^t \psi_j(u) S_{hj}(u) \left(\zeta_j^K(s, du) - \zeta_j(s, du) \right),$

Let $\boldsymbol{\eta}_j^{K*}(s, t) = \int_0^t \psi_j(u) S_{hj}(u) \left(\zeta_j^K(s, du) - \zeta_j(s, du) \right).$

We consider the $(h, k)^{th}$ component of the predictable covariation process for (A.9) in two parts. The first term is $n \int_0^\infty \boldsymbol{\eta}_j^K(s, t)^T \mathbf{I}_\lambda(s) \boldsymbol{\eta}_j^K(s, t) ds$, where $\mathbf{I}_\lambda(s)$ is as defined in (A.7). Performing the multiplication produces

$$n \int_0^\infty \sum_{l=i}^K \sum_{i=1}^{n_l} \eta_{lij}^K(s, t)^2 I[T_{lij} \geq t] \lambda_{li}^c(s) ds.$$

We assume that for all i and l

- $\eta_{lij}^{K*}(s, t)^2 \lambda_{li}^c(s)$ is integrable
- $n \int_0^\infty \eta_{lij}^K(s, t)^2 I[T_{lij} \geq t] \lambda_{li}^c(s) ds \xrightarrow{p} \int_0^\infty \eta_{lij}^{K*}(s, t)^2 I[T_{lij} \geq t] \lambda_{li}^c(s) ds,$

where the η_{lij}^{K*} are the components of $\boldsymbol{\eta}_j^{K*}$.

For the second term of (A.9), we assume the existence, finiteness, and non-singularity of the matrix $\mathbf{a}(s) = E \left[\widetilde{\mathbf{X}}_{hi}(s) \widetilde{\mathbf{X}}_{hi}^T(s) \right]$. Then

$$\begin{aligned} & \sqrt{n} \int_0^\infty \boldsymbol{\eta}_j^K(s, t) \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s) d\mathbf{M}^c(s) \\ &= \frac{1}{\sqrt{n}} \int_0^\infty \boldsymbol{\eta}_j^K(s, t) \widetilde{\mathbf{X}}(s) (n^{-1} \mathbf{A}(s))^{-1} \widetilde{\mathbf{X}}^T(s) d\mathbf{M}^c(s) \\ &= \frac{1}{\sqrt{n}} \int_0^\infty \boldsymbol{\eta}_j^K(s, t) \widetilde{\mathbf{X}}(s) \mathbf{a}^{-1}(s) \widetilde{\mathbf{X}}^T(s) d\mathbf{M}^c(s) + o_p(n^{-1/2}) \\ &= \frac{1}{\sqrt{n}} \int_0^\infty \boldsymbol{\eta}_j^K(s, t) \mathbf{P}(s) d\mathbf{M}^c(s) + o_p(n^{-1/2}), \end{aligned}$$

where $\mathbf{P}(s) = p \lim_{n \rightarrow \infty} \widetilde{\mathbf{X}}(s) \mathbf{a}^{-1}(s) \widetilde{\mathbf{X}}^T(s)$, which we assume to exist. The second term of the $(h, k)^{th}$ component of the predictable covariation process for (A.9) is then

$$\frac{1}{n} \int_0^\infty \boldsymbol{\eta}_j^K(s, t) \mathbf{P}(s) \mathbf{I}_\lambda(s) \mathbf{P}(s) \boldsymbol{\eta}_j^K(s, t) ds.$$

We now assume that

$$\frac{1}{n} \int_0^\infty \boldsymbol{\eta}_j^K(s, t) \mathbf{P}(s) \mathbf{I}_\lambda(s) \mathbf{P}(s) \boldsymbol{\eta}_j^K(s, t) ds \xrightarrow{p} \int_0^\infty \boldsymbol{\eta}_j^{K*}(s, t) \mathbf{P}(s) \mathbf{I}_\lambda(s) \mathbf{P}(s) \boldsymbol{\eta}_j^{K*}(s, t) ds$$

Collecting our results thus far gives us that

$$\langle \widehat{Z}_{hj}(t), \widehat{Z}_{kj}(t) \rangle \xrightarrow{p} \int_0^t \psi_j^2(s) y_{hj}^*(s) y_{kj}^*(s) (y_j^*(s))^{-1} \lambda_j(s) ds +$$

$$\int_0^\infty \boldsymbol{\eta}_j^{K*}(s, t)^T \mathbf{I}_\lambda(s) \boldsymbol{\eta}_j^{K*}(s, t) ds - \int_0^\infty \boldsymbol{\eta}_j^{K*}(s, t)^T \mathbf{P}(s) \mathbf{I}_\lambda(s) \mathbf{P}(s) \boldsymbol{\eta}_j^{K*}(s, t) ds,$$

(i.e.) that the predictable covariation process of $\widehat{\mathbf{Z}}_j(t)$ approaches a deterministic limit.

We now make a general assumption to satisfy the Lindeberg condition of the martingale central limit theorem for the second term (A.9) (the Lindeberg condition for (A.8) has been addressed). Let $\mathbf{H}_j^K(s, t) = \sqrt{n} \boldsymbol{\eta}_j^K(s, t)^T [\mathbf{I} - \widetilde{\mathbf{X}}(s) \mathbf{A}^{-1}(s) \widetilde{\mathbf{X}}^T(s)]$, and note that the second term of (A.6) can be expressed as

$$\int_0^\infty \mathbf{H}_j^K(s, t) d\mathbf{M}^c(s) = \sum_{k=1}^K \sum_{i=1}^{n_k} \int_0^\infty H_{kij}^K(s, t) dM_{ki}^c(s),$$

where $H_{kij}^K(s, t)$ are the lexicographically ordered components of $\mathbf{H}_j^K(s, t)$. We then assume (cf. eq. 2.5.8, [6]) that

$$\sum_{k=1}^K \sum_{i=1}^{n_k} \int_0^\infty H_{kij}^K(s, t)^2 I(|H_{kij}^K(s, t)| > \epsilon) \lambda_{ki}^c(s) ds \xrightarrow{p} 0,$$

(i.e.) we assume that the second term (A.9) of the h^{th} component of $\widehat{\mathbf{Z}}_j(t)$ approaches a continuous limit. This assumption, coupled with convergence of the predictable covariation process to a deterministic limit, guarantees the convergence of (A.9) to a normal limit. Putting the weak convergence of (A.8) and (A.9) while recognizing their orthogonality proves Theorem III.1. ■

Proof of Lemma IV.1.

We begin by “vectorizing” the results from the proof of Theorem II.1. Specifically, Theorem II.1 asserts that $\widehat{N}_{ij}(t) = N_{ij}^*(t) - N_{ij}^*(t) \int_0^{U_{ij}^*} \widehat{K}_i(s)^{-1} d\widehat{M}_i^c(s)$ and $\widehat{Y}_{ij}(t) = Y_{ij}^*(t) - Y_{ij}^*(t) \int_0^{T_{ij}^*+t} \widehat{K}_i(s)^{-1} d\widehat{M}_i^c(s)$. The corresponding vector/matrix results are

$$\widehat{\mathbf{N}}_j(t) - \mathbf{N}_j^*(t) = - \left[\int_0^\infty \mathbf{diag} \left\{ \frac{I[s < U_{ij}^*]}{\widehat{K}_i(s)} \right\} d\widehat{M}_i^c(s) \right] \mathbf{N}_j^*(t), \quad (\text{A.10})$$

$$\begin{aligned} \widehat{\mathbf{Y}}_j(t) - \mathbf{Y}_j^*(t) &= \mathbf{diag} \{ \widehat{Y}_{ij}(t) \} \mathbf{X}(\mathbf{T}_j + t) - \mathbf{diag} \{ Y_{ij}^*(t) \} \mathbf{X}(\mathbf{T}_j^* + t) \\ &= \mathbf{diag} \{ \widehat{Y}_{ij}(t) \} \mathbf{X}(\mathbf{T}_j + t) - \mathbf{diag} \{ Y_{ij}^*(t) \} \mathbf{X}(\mathbf{T}_j^* + t) + \\ &\quad \mathbf{diag} \{ \widehat{Y}_{ij}(t) \} \mathbf{X}(\mathbf{T}_j^* + t) - \mathbf{diag} \{ \widehat{Y}_{ij}(t) \} \mathbf{X}(\mathbf{T}_j^* + t) \end{aligned}$$

$$\begin{aligned}
&= \mathbf{diag} \left\{ \widehat{Y}_{ij}(t) - Y_{ij}^*(t) \right\} \mathbf{X}(\mathbf{T}_j^* + t) + \\
&\quad \mathbf{diag} \left\{ \widehat{Y}_{ij}(t) \right\} \left(\mathbf{X}(\mathbf{T}_j + t) - \mathbf{X}(\mathbf{T}_j^* + t) \right) \\
&= \mathbf{diag} \left\{ \widehat{Y}_{ij}(t) - Y_{ij}^*(t) \right\} \mathbf{X}(\mathbf{T}_j^* + t) \\
&= -\mathbf{diag} \left\{ Y_{ij}^*(t) \int_0^\infty \frac{I[s < T_{ij}^* + t]}{\widehat{K}_i(s)} d\widehat{M}_i^c(s) \right\} \mathbf{X}(\mathbf{T}_j^* + t). \\
&= -\mathbf{diag} \left\{ \int_0^\infty \frac{I[s < T_{ij}^* + t]}{\widehat{K}_i(s)} d\widehat{M}_i^c(s) \right\} \mathbf{Y}_j^*(t) \tag{A.11}
\end{aligned}$$

We note that the fourth line of the derivation of the result for $\widehat{\mathbf{Y}}_j(t) - \mathbf{Y}_j^*(t)$ asserts that $\mathbf{diag} \left\{ \widehat{Y}_{ij}(t) \right\} \left(\mathbf{X}(\mathbf{T}_j + t) - \mathbf{X}(\mathbf{T}_j^* + t) \right)$ is a zero matrix; we briefly illustrate this. The $(i, k)^{th}$ element of this matrix expression is $\widehat{Y}_{ij}(t) \left(X_{ik}(T_{ij} + t) - X_{ik}(T_{ij}^* + t) \right)$. For $T_{ij}^* \leq C_i$, we note that $T_{ij}^* = T_{ij}$ and hence $X_{ik}(T_{ij}^* + t) = X_{ik}(T_{ij} + t)$ and the $(i, k)^{th}$ element is zero. For $T_{ij}^* > C_i$, $\widehat{Y}_{ij}(t) = I[U_{ij} - T_{ij} \geq t, T_{ij} \leq C_i] / K_i(T_{ij} + t) = 0$ and again the $(i, k)^{th}$ element is zero. Therefore, $\mathbf{diag} \left\{ \widehat{Y}_{ij}(t) \right\} \left(\mathbf{X}(\mathbf{T}_j + t) - \mathbf{X}(\mathbf{T}_j^* + t) \right) = \mathbf{0}$.

Both (A.10) and (A.11) involve stochastic integration with respect to the function $d\widehat{M}_i^c(s)$. The equality $d\widehat{M}_i^c(u) = dM_i^c(u) - I[T_i \geq u] d[\widehat{\Lambda}_i^c - \Lambda_i^c](u)$ noted in the proof of Lemma II.4 takes vector form $d\widehat{\mathbf{M}}^c(u) = d\mathbf{M}^c(u) - \mathbf{diag} \{ I[T_i \geq u] \} d[\widehat{\Lambda}^c - \Lambda^c](u)$. By Aalen's linear model for the censoring hazard, we have that

$$\widehat{\Lambda}^c(u) - \Lambda^c(u) = \int_0^u \mathbf{Y}^c(r) \left(\mathbf{Y}^c(r)^T \mathbf{Y}^c(r) \right)^{-1} \mathbf{Y}^c(r)^T d\mathbf{M}^c(r) + o_p(n^{-1/2}),$$

where $\mathbf{Y}^c(r)$ is as defined in the statement of the theorem. Therefore, we have that $d\widehat{\mathbf{M}}^c(u) = \mathbf{P}(u)d\mathbf{M}^c(u)$, where $\mathbf{P}(u)$ is as in the statement of the theorem, and (A.10) and (A.11) become

$$\begin{aligned}
\widehat{\mathbf{N}}_j(t) &= \left(\mathbf{I} - \int_0^\infty \mathbf{diag} \left\{ \frac{I[s < U_{ij}^*]}{\widehat{K}_i(s)} dM_i^c(s) \right\} \mathbf{P}(s) \right) \mathbf{N}_j^*(t) \\
&= \left(\mathbf{I} - \int_0^\infty \mathbf{diag} \left\{ \frac{I[s < U_{ij}^*]}{K_i(s)} dM_i^c(s) \right\} \mathbf{P}(s) \right) \mathbf{N}_j^*(t) + o_p(n^{-1/2}) \\
&= (\mathbf{I} - \mathbf{D}_N(s)\mathbf{P}(s)) \mathbf{N}_j^*(t) + o_p(n^{-1/2}) \tag{A.12}
\end{aligned}$$

$$\begin{aligned}
\widehat{\mathbf{Y}}_j(t) &= \left(\mathbf{I} - \int_0^\infty \mathbf{diag} \left\{ \frac{I[s < T_{ij}^* + t]}{\widehat{K}_i(s)} dM_i^c(s) \right\} \mathbf{P}(s) \right) \mathbf{Y}_j^*(t) \\
&= \left(\mathbf{I} - \int_0^\infty \mathbf{diag} \left\{ \frac{I[s < T_{ij}^* + t]}{K_i(s)} dM_i^c(s) \right\} \mathbf{P}(s) \right) \mathbf{Y}_j^*(t) + o_p(n^{-1/2}) \\
&= (\mathbf{I} - \mathbf{D}_Y(s, t)\mathbf{P}(s)) \mathbf{Y}_j^*(t) + o_p(n^{-1/2}), \tag{A.13}
\end{aligned}$$

where $\mathbf{D}_N(s)$ and $\mathbf{D}_Y(s, t)$ are as in the statement of the theorem. We briefly note here the product commutativity of symmetric matrices (such as $\mathbf{P}(u)$) and diagonal matrices, which will be used below. From (A.13), we note that a generalized inverse of $\widehat{\mathbf{Y}}_j^-(t)$ is given by

$$\widehat{\mathbf{Y}}_j^-(t) = \mathbf{Y}_j^{*-}(t) (\mathbf{I} - \mathbf{D}_Y(s, t)\mathbf{P}(u))^{-},$$

and that a useful generalized inverse of $\mathbf{I} - \mathbf{D}_Y(s, t)\mathbf{P}(u)$ is given by the matrix $(\mathbf{I} + (\mathbf{I} - \mathbf{D}_Y(s, t)\mathbf{P}(u))^{-}\mathbf{D}_Y(s, t)\mathbf{P}(u))$.

We now note that based on (A.12) and (A.13),

$$\begin{aligned} \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) &= \int_0^t \widehat{\mathbf{Y}}_j^-(s) d\widehat{\mathbf{N}}_j(s) - \mathbf{B}_j(t) \\ &= \int_0^t \mathbf{Y}_j^{*-}(s) \left(\mathbf{I} + (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-}\mathbf{D}_Y(u, s)\mathbf{P}(u) \right) \times \\ &\quad (\mathbf{I} - \mathbf{D}_N(u)\mathbf{P}(u)) d\mathbf{N}_j^*(s) - \mathbf{B}_j(t) + o_p(n^{-1/2}). \end{aligned}$$

Expanding this expression produces

$$\begin{aligned} \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) &= \int_0^t \mathbf{Y}_j^{*-}(s) d\mathbf{N}_j^*(s) - \mathbf{B}_j(t) + \\ &\quad \int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-}\mathbf{D}_Y(u, s)\mathbf{P}(u) d\mathbf{N}_j^*(s) - \\ &\quad \int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{D}_N(u)\mathbf{P}(u) + (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-}\mathbf{D}_Y(u, s)\mathbf{P}(u) \times \\ &\quad \mathbf{D}_N(u)\mathbf{P}(u)) d\mathbf{N}_j^*(s) + o_p(n^{-1/2}). \end{aligned} \tag{A.14}$$

Label the three summands of (A.14) as I, II, and III. By standard martingale methods (see Eq. 7.4.7 of Andersen, et al [6]), the first term satisfies

$$\text{I} = \int_0^t J_j^*(s) \mathbf{Y}_j^{*-}(s) d\mathbf{M}_j^*(s) + o_p(n^{-1/2}). \tag{A.15}$$

For term II, we note that the diagonal matrix $\mathbf{D}_Y(u, s)$ and the symmetric matrix $\mathbf{P}(u)$ commute, and that we can interchange the implied order of integration to produce

$$\begin{aligned} \text{II} &= \int_0^\infty \int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-} \mathbf{diag} \left\{ \frac{I[u < T_{ij}^* + s]}{K_i(s)} dN_{ij}^*(s) \right\} \mathbf{P}(u) d\mathbf{M}^c(u) + \\ &\quad o_p(n^{-1/2}) \end{aligned} \tag{A.16}$$

Factoring the third term produces

$$\text{III} = \int_0^t \mathbf{Y}_j^{*-}(s) \left(\mathbf{I} + (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1} \mathbf{D}_Y(u, s)\mathbf{P}(u) \right) \mathbf{D}_N(u)\mathbf{P}(u) d\mathbf{N}_j^*(s) + o_p(n^{-1/2}),$$

which, after interchanging the order of integration implied by $\mathbf{D}_N(u)$ and $d\mathbf{N}_j^*(s)$ and substituting in a simplified version of $(\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1}$, becomes

$$\begin{aligned} \text{III} &= \int_0^\infty \int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1} \text{diag} \left\{ \frac{I[u < U_{ij}^*]}{K_i(u)} d\mathbf{N}_{ij}^*(s) \right\} \mathbf{P}(u) d\mathbf{M}^c(u), \\ &+ o_p(n^{-1/2}) \end{aligned} \quad (\text{A.17})$$

Putting (A.16) and (A.17) together produces

$$\begin{aligned} \text{II} - \text{III} &= \int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1} \text{diag} \left\{ \left(\frac{I[u < U_{ij}^*]}{K_i(u)} - \frac{I[u < T_{ij}^* + s]}{K_i(u)} \right) d\mathbf{N}_{ij}^*(s) \right\} \right] \mathbf{P}(u) d\mathbf{M}^c(u) + o_p(n^{-1/2}), \\ &= \int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1} \text{diag}\{\mathbf{Q}(u, s)\} \text{diag}\{d\mathbf{N}_j^*(s)\} \right] \times \\ &\quad \mathbf{P}(u) d\mathbf{M}^c(u) + o_p(n^{-1/2}), \end{aligned} \quad (\text{A.18})$$

Adding (A.15) to (A.18) produces

$$\begin{aligned} \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) &= \int_0^t J_j^*(s) \mathbf{Y}_j^{*-}(s) d\mathbf{M}^*(s) + \\ &\quad \int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1} \text{diag}\{\mathbf{Q}(u, s)\} \text{diag}\{d\mathbf{N}_j^*(s)\} \right] \times \\ &\quad \mathbf{P}(u) d\mathbf{M}^c(u) + o_p(n^{-1/2}), \end{aligned} \quad (\text{A.19})$$

and the lemma is proven. ■

Proof of Theorem IV.2.

We recall the martingale representation of $\widehat{\mathbf{B}}_j(t)$ given above (A.19) and its predictable variation process:

$$\begin{aligned} \langle \widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t) \rangle &= \int_0^t J_j^*(s) \mathbf{Y}_j^{*-}(s) \text{diag}\{\boldsymbol{\lambda}_j(s)\} \mathbf{Y}_j^{*-}(s)^\top ds + \\ &\quad \int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s)\mathbf{P}(u))^{-1} \text{diag}\{\mathbf{Q}(u, s)\} \text{diag}\{d\mathbf{N}_j^*(s)\} \right] \times \end{aligned}$$

$$\begin{aligned} & \mathbf{P}(u) \mathbf{diag}\{\boldsymbol{\lambda}^c(u)\} \mathbf{P}(u)^T \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag}\{\mathbf{Q}(u, s)\} \right. \\ & \left. \mathbf{diag}\{d\mathbf{N}_j^*(s)\} \right]^T du, \end{aligned} \quad (\text{A.20})$$

As with the proof of Theorem III.1, we consider the two martingales composing $\widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t)$ separately given their orthogonality. We first note that the first term of (A.19) directly corresponds to the martingale representation of Aalen's linear model for counting processes as described in Andersen, et al [6] (cf. Equation 7.4.7, pg. 564). The conditions required for the weak convergence of this martingale are defined therein, which we state here:

1. the existence of continuous functions $y_{j,k}^{(1)}(s)$, $y_{j,kl}^{(2)}(s)$, and $y_{j,klm}^{(3)}(s)$, such that for all $0 \leq k, l, m \leq p$ as $n \rightarrow \infty$

- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{j,ik}^*(s) - y_{j,k}^{(1)}(s) \right| \xrightarrow{p} 0$
- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{j,ik}^*(s) Y_{j,il}^*(s) - y_{j,kl}^{(2)}(s) \right| \xrightarrow{p} 0$
- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{j,ik}^*(s) Y_{j,il}^*(s) Y_{j,im}^*(s) - y_{j,klm}^{(3)}(s) \right| \xrightarrow{p} 0,$

where $Y_{j,kl}^*$ denotes the $(k, l)^{th}$ element of the matrix \mathbf{Y}_j^* ;

2. for all $l = 0, \dots, p,$

$$n^{-1/2} \sup_{\substack{s \in [0, t] \\ k=1, \dots, n}} |Y_{j,kl}^*(s)| \xrightarrow{p} 0;$$

3. the matrix $\mathbf{y}_j^{(2)}(s) = \{y_{j,kl}^{(2)}(s)\}$ is non-singular.

Under these conditions, the martingale $\sqrt{n} \int_0^t J_j^*(s) \mathbf{Y}_j^{*-}(s) d\mathbf{M}_j^*(s)$ defined in the first term of (A.19) converges to a normal limit with covariance matrix defined by the first integral of (A.20). The details of the proof can be found in the proof of Theorem VII.4.1 in Andersen, et al [6]. Note that although these conditions look somewhat esoteric, they directly correspond to the conditions necessary for the application of the martingale central limit theorem. Condition 1 establishes a law of large number for the products of the matrix $\mathbf{Y}_j^*(s)$ that form the least squares generalized inverse $(\mathbf{Y}_j^*(s)^T \mathbf{Y}_j^*(s))^{-1} \mathbf{Y}_j^*(s)^T$, guaranteeing their convergence to a continuous, deterministic limit which in turn guarantees the convergence of the predictable covariation process to a continuous, deterministic limit. Condition 2 is a Lindeberg condition on the elements of the covariate matrix $\mathbf{Y}_j^*(s)$, and condition 3 guarantees that $\mathbf{Y}_j^*(s)^T \mathbf{Y}_j^*(s)$ converges to an invertible limit. Additional

assumptions on the covariates generally lead to more explicit conditions for weak convergence, such as assuming that the counting (N_{ij}^*, Y_{ij}^*) and covariate processes (\mathbf{X}_i) are together i.i.d. (see Andersen, et al [6], Ex. VIII.4.4 for further details).

We now turn attention to the second term of (A.19),

$$\sqrt{n} \int_0^t \int_0^\infty \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag}\{\mathbf{Q}(u, s)\} \mathbf{diag}\{d\mathbf{N}_j^*(s)\} \right] \mathbf{P}(u) d\mathbf{M}^c(u). \quad (\text{A.21})$$

For notational convenience, let $\mathbf{R}(u, s) = (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))$. By assumptions 1 and 3 on the least squares generalized inverse of $\mathbf{Y}_j^*(s)$ from the previous page, we note that

$$\begin{aligned} & \int_0^t \mathbf{Y}_j^{*-}(s) \mathbf{R}(u, s)^{-1} \mathbf{diag}\{\mathbf{Q}(u, s)\} \mathbf{diag}\{d\mathbf{N}_j^*(s)\} \\ &= \frac{1}{n} \int_0^t \left(n^{-1} \mathbf{Y}_j(s)^T \mathbf{Y}_j(s) \right)^{-1} \mathbf{Y}_j(s)^T \mathbf{R}(u, s)^{-1} \mathbf{diag}\{\mathbf{Q}(u, s)\} \mathbf{diag}\{d\mathbf{N}_j^*(s)\} \\ &= \frac{1}{n} \int_0^t \mathbf{y}_j^{(2)}(s)^{-1} \left[\mathbf{Y}_j(s)^T \mathbf{R}(u, s)^{-1} \mathbf{diag}\{\mathbf{Q}(u, s)\} \mathbf{diag}\{d\mathbf{N}_j^*(s)\} \right] + o_p(n^{-1/2}) \end{aligned}$$

where $\mathbf{y}_j^{(2)}(s)$ is the matrix defined in condition 1 on the previous page. Performing the stochastic integration implied above produces the $(p+1) \times n$ matrix with $(k, l)^{th}$ element

$$\frac{1}{n} \sum_{h=1}^n \sum_{i=0}^p Q_l(u, U_{ij}^* - T_{ij}^*) Y_{hi}(U_{ij}^* - T_{ij}^*) R_{hl}^-(u, U_{ij}^* - T_{ij}^*) y_{j,ki}^{(2)}(U_{ij}^* - T_{ij}^*)^{-1}, \quad (\text{A.22})$$

where $Q_l(u, s)$ is the l^{th} element of the vector $\mathbf{Q}(u, s)$, $R_{hl}^-(u, s)$ the $(h, l)^{th}$ element of the generalized inverse of $\mathbf{R}(u, s)$, and $y_{j,ki}^{(2)}(s)$ the $(k, i)^{th}$ element of $\mathbf{y}_j^{(2)}$. We now assume the existence, finiteness, and continuity (in u) of the quantity

$$\sum_{i=0}^p y_{j,ki}^{(2)}(U_{ij}^* - T_{ij}^*)^{-1} E \left[Q_l(u, U_{ij}^* - T_{ij}^*) Y_{hi}(U_{ij}^* - T_{ij}^*) R_{hl}^-(u, U_{ij}^* - T_{ij}^*) \right],$$

to which (A.22) converges via the law of large numbers. In summary, we have that the inner integral of (A.21) converges to a $(p+1) \times n$ matrix which we denote $\boldsymbol{\xi}(u, t)$.

We now turn attention to the matrix $\mathbf{P}(u)$. Since $\mathbf{P}(u)$ is a matrix product arising from the estimation of the censoring hazard via Aalen's linear model, we can impose the technical conditions previously used for $\mathbf{Y}_j^*(s)$ onto $\mathbf{Y}^c(s)$:

1. the existence of continuous functions $y_h^{(c1)}(s)$, $y_{hk}^{(c2)}(s)$, $y_{hkl}^{(c3)}(s)$, and $y_{hklm}^{(c4)}(s)$ such that for all $0 \leq h, k, l, m \leq p$ as $n \rightarrow \infty$

- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{ih}^c(s) - y_h^{(c1)}(s) \right| \xrightarrow{p} 0$
- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{ih}^c(s) Y_{ih}^c(s) - y_{hk}^{(c2)}(s) \right| \xrightarrow{p} 0$
- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{ih}^c(s) Y_{ik}^c(s) Y_{il}^c(s) - y_{hkl}^{(c3)}(s) \right| \xrightarrow{p} 0,$
- $\sup_{s \in [0, t]} \left| \frac{1}{n} \sum_{i=1}^n Y_{ih}^c(s) Y_{ik}^c(s) Y_{il}^c(s) Y_{im}^c(s) - y_{hklm}^{(c4)}(s) \right| \xrightarrow{p} 0,$

where Y_{kl}^c denotes the $(k, l)^{th}$ element of the matrix \mathbf{Y}^c ;

2. for all $l = 0, \dots, p,$

$$n^{-1/2} \sup_{\substack{s \in [0, t] \\ k=1, \dots, n}} |Y_{kl}^c(s)| \xrightarrow{p} 0;$$

3. the matrix $\mathbf{y}^{(c2)}(s) = \{y_{kl}^{(c2)}(s)\}$ is non-singular.

We note that the fourth assumption for condition 1 is required since $\mathbf{P}(u)$ involves a 4-product of the matrix $\mathbf{Y}_c(u)$. These conditions ensure that $\mathbf{P}(u)$ approaches a continuous, deterministic limit denoted $\mathbf{p}(u)$ as follows:

$$\begin{aligned} \mathbf{P}(u) &= \mathbf{I} - \mathbf{Y}_c(u) \left(\mathbf{Y}_c(u)^T \mathbf{Y}_c(u) \right)^{-1} \mathbf{Y}_c(u)^T \\ &= \mathbf{I} - \frac{1}{n} \mathbf{Y}_c(u) \left(n^{-1} \mathbf{Y}_c(u)^T \mathbf{Y}_c(u) \right)^{-1} \mathbf{Y}_c(u)^T \\ &= \mathbf{I} - \frac{1}{n} \mathbf{Y}_c(u) \mathbf{y}^{(c2)}(u)^{-1} \mathbf{Y}_c(u)^T + o_p(n^{-1/2}) \\ &= \mathbf{I} - \mathbf{y}^{(c1)}(u) \mathbf{y}^{(c2)}(u)^{-1} \mathbf{y}^{(c1)}(u)^T + o_p(n^{-1/2}) \\ &= \mathbf{p}(u) + o_p(n^{-1/2}). \end{aligned}$$

We then have that the second term of the martingale representation (A.21) is

$$\int_0^\infty \boldsymbol{\xi}(u) \mathbf{p}(u) d\mathbf{M}^c(u),$$

where we note that $\boldsymbol{\xi}(u) \mathbf{p}(u)$, being continuous and deterministic, is predictable and hence its stochastic integral with respect to the zero-mean martingale $d\mathbf{M}^c(u)$ is also a zero-mean martingale with variance given by the continuous, deterministic predictable covariation process

$$\int_0^t \boldsymbol{\xi}(u, t) \mathbf{p}(u) \mathbf{diag}\{\boldsymbol{\lambda}^c(u)\} \mathbf{p}^T(u) \boldsymbol{\xi}^T(u, t) du, \quad (\text{A.23})$$

to which the predictable covariation process (A.20) converges.

We make a general assumption to satisfy the Lindeberg condition required for

asymptotic normality. Let $\mathbf{H}(u, t)$ denote the $(p + 1) \times n$ matrix

$$\mathbf{H}(u, t) = \sqrt{n} \left[\int_0^t \mathbf{Y}_j^{*-}(s) (\mathbf{I} - \mathbf{D}_Y(u, s) \mathbf{P}(u))^{-1} \mathbf{diag}\{\mathbf{Q}(u, s)\} \mathbf{diag}\{d\mathbf{N}_j^*(s)\} \right] \mathbf{P}(u).$$

Then the second term of the martingale representation for $\widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t)$ in (A.21) can be expressed as $\int_0^\infty \mathbf{H}(u, t) d\mathbf{M}^c(u)$. The k^{th} term of this vector is

$$\sum_{i=1}^n \int_0^\infty H_{ki}(u, t) dM_i^c(u).$$

We then assume (cf. eq. 2.5.8, [6]) that

$$\sum_{i=1}^n \int_0^\infty H_{ki}^2(u, t) I[|H_{ki}(u, t)| > \epsilon] \lambda_{ki}^c(s) ds \xrightarrow{p} 0, \quad (\text{A.24})$$

for all k , (i.e.) we assume that each of the $p + 1$ components of the second term (A.21) approaches a continuous limit (that the ‘‘jumps’’ of the associated martingale are asymptotically arbitrarily small). The convergence of the predictable covariation process (A.20) to a continuous, deterministic limit (A.23) coupled with Lindeberg condition (A.24) ensures that the second term (A.21) in the martingale expansion of $\widehat{\mathbf{B}}_j(t) - \mathbf{B}_j(t)$ converges weakly to a normal limit. This result, coupled with the weak convergence of the first term of (A.19) to a normal limit, while recognizing the orthogonality of the two terms of (A.19), proves Theorem IV.2. ■

CURRICULUM VITAE

- NAME: Douglas J. Lorenz
- ADDRESS: Department of Mathematics
University of Louisville
Louisville, KY 40292
- EDUCATION: B.S. Mathematics, University of Louisville, 1999
M.A. Mathematics, University of Louisville, 2001
M.S.P.H. Biostatistics/Decision Science,
University of Louisville, 2002
Ph.D. Applied and Industrial Mathematics
University of Louisville, 2011
- PROFESSIONAL SOCIETIES: American Statistical Association
- CERTIFICATIONS: PStat, Accredited Statistician of the American Statistical Association
- AWARDS: Graduate Dean's Citation, University of Louisville, 2002
Hohman Fellowship, Department of Mathematics,
University of Louisville, 2006
- PUBLICATIONS: **Lorenz DJ**, Datta S, Harkema SJ. Marginal association measures for clustered data. Submitted.
- Lorenz DJ**, Datta S. A log-rank test for waiting times from a multi-stage model. Submitted.
- Lorenz DJ**, Datta S, Harkema SJ. Longitudinal patterns of functional recovery in patients with incomplete spinal cord injury receiving activity-based rehabilitation. Preprint.
- Datta S, **Lorenz DJ**, Harkema SJ. A dynamic longitudinal evaluation of the utility of the Berg Balance Scale in patients with motor incomplete spinal cord injury. Preprint.
- Forrest GF, **Lorenz DJ**, Hutchinson K, Van Hiel L, Basso DM, Datta S, Sisto SA, Harkema SJ. Relationships between balance and walking measures at baseline and after

locomotor training in incomplete SCI: impact of functional recovery. Preprint.

Harkema SJ, Schmidt-Read M, **Lorenz D**, Edgerton VR, Behrman A. Balance and ambulation improvements in individuals with chronic incomplete spinal cord injury using Locomotor Training-based rehabilitation. *Arch Phys Med Rehabil*. Accepted.

Tate SB, Blackwell L, **Lorenz DJ**, Steptoe MM, Culligan PJ. Randomized Trial of Fascia Lata and Polypropylene Mesh for Abdominal Sacrocolpopexy: 5-Year Follow-up. *Int Urogynecol J Pelvic Floor Dysfunct*. 2010 Aug 27.

Jackson JM, Fowler JF, Callen, JP, **Lorenz DJ**. Mycophenolate Mofetil for the Treatment of Chronic Dermatitis: An Open-Label Study of 16 Patients. *J Drugs Dermatol*. 2010 Apr; 9(4): 356-62

Condra CS, Parbhu D, **Lorenz D**, Herr SM. Charges and complications associated with the medical evaluation of febrile infants. *Pediatr Emerg Care*. 2010 Mar; 26(3): 186-91.

Alur AA, Rane MJ, Scheetz JP, **Lorenz DJ**, Gettleman R. Simulated Microbe Removal Around Finger Rings Using Different Hand Sanitation Methods. *International Journal of Oral Science*, 1(3): 138-44, 2009.

Pierce MC, Kaczor K, Aldridge S, O'Flynn J, **Lorenz DJ**. Bruising Characteristics Discriminate Physical Child Abuse From Accidental Trauma in Young Children. *Pediatrics*. 2010 Jan; 125(1):67-74.

Padmanabhan P, Berkenbosch JW, **Lorenz DJ**, Pierce MC. Evaluation of Cerebral Oxygenation during Procedural Sedation in Children using Near Infrared Spectroscopy. *Ann Emerg Med*. Aug 2009; 54(2).

Datta S, **Lorenz DJ**, Morrison S, Ardolino E, Harkema SJ. A Multivariate Examination of Temporal Changes in Berg Balance Scale Items for Patients with ASIA Impairment Scale C and D Spinal Cord Injuries. *Arch Phys Med Rehabil*. 2009 Jul; 90(7): 1208-17.

Fortu JM, Kim IK, Cooper A, Condra C, **Lorenz DJ**, Pierce MC. Psychiatric patients in the pediatric emergency department undergoing routine urine toxicology screens for

medical clearance: results and use. *Pediatr Emerg Care*. 2009 Jun; 25(6): 387-92.

Spiller HA, **Lorenz DJ**. Trends in Volatile Substance Abuse (VSA). *J Add Dis*, 28:130-136. 2009.

Spiller H, **Lorenz DJ**, Bailey EJ, Dart RC. Epidemiological trends in abuse and misuse of prescription opioids. *J Add Dis*, 2009;28(2):130-6.

Ruble LI, McDuffie AS, King AS, **Lorenz DJ**. Caregiver responsiveness and social interaction behaviors of young children with autism. *Topics in Early Childhood Special Education*. Nov 2008; 28: 158-70.

Warkentine F, Pierce MC, Kim IK, **Lorenz DJ**, Paul RI. The anatomic relationship of femoral vein to femoral artery in euvoletic pediatric patients by ultrasonography: implications for pediatric femoral central venous access. *Acad Emerg Med* 2008 May 15; 15(5):426-30.

Helm, CW, Helm CW, **Lorenz DJ**, Meyer NJ, Rising WR, Wilff JL. Retinoids for preventing the progression of cervical intra-epithelial neoplasia. *Cochrane Database Syst Rev*, 2007 Oct 17; (4):CD003296. Review.

Koller DM, Myers AB, **Lorenz D**, Godambe, SA. Effectiveness of Oxycodone, Ibuprofen, or the Combination in the Initial Management of Orthopedic Injury-Related Pain in Children. *Ped Emerg Care*, Sep 2007; 23(9).

Hoffman JS, Cunningham DR, **Lorenz DJ**. Auditory thresholds and factors contributing to hearing loss in a large sample of percussionists. *Medical Problems of Performing Artists*, June 2006; 21(2): 47-58.

Raj AB, Condurache T, Bertolone S, Williams D, **Lorenz D**, Sobczyk W. Qualitative assessment of ventricular function in sickle cell disease: effect of long term erythrocytapheresis. *Pediatr Blood Cancer*, Dec. 2005; 45(7): 976-981.

Chao C, Studts JL, Abell T, Hadley T, Roetzer L, Dineen S, **Lorenz D**, YoussefAgha A, McMasters KM. Adjuvant chemotherapy for breast cancer: how presentation of recurrence risk influences decision-making. *J Clin Oncol*, 2003 Dec 1; 21(23): 4299-305.

NATIONAL
PRESENTATIONS: Datta S, **Lorenz D**, Schmidt-Read M, Ardolino E,

Morrison S, Harkema S. A multivariate examination of temporal changes in Berg Balance Scale variables for patients with ASIA C and D spinal cord injuries. 33rd Annual Meeting of the American Spinal Injury Association, Tampa, FL, May 2007. **Winner, Best Poster Award** .

Lajoie AS, De La Cruz R, **Lorenz DJ**, Ramirez JA. Predicting time to clinical stability and length of hospital stay in patients with community-acquired pneumonia. Joint Statistical Meetings, Minneapolis, MN, August 2005.

Jackson JM, Fowler JF, Callen JP, **Lorenz D**. Mycophenolate mofetil for the treatment of chronic dermatitis: An open-label study of 16 patients. [Abstract], J Am Acad Dermatol, March 2005; 52(3, S1): P78.

Kurian RT, Barve SS, Ritchie CS, Stagner IJ, Adams KA, **Lorenz DJ**, Mokshagundam SL. Resistance Training Improves Neurological Function and Cardiovascular Risk Factors in Elderly Subjects with Type 2 Diabetes Mellitus and Peripheral Neuropathic Symptoms. 64th Annual Scientific Sessions of the American Diabetes Association, Orlando, FL, June 2004.

Fowler J, **Lorenz D**, Parneix-Spake A. Evaluation of skin reactions to pimecrolimus cream 1% in patients allergic to propylene glycol. American Academy of Dermatology 63rd Annual Meeting, New Orleans, LA, February 2004.

**FUNDED
RESEARCH:**

Harkema SJ (PI). Development of neural recovery rehab & research centers. Christopher Reeve Paralysis Foundation. RFA7-20-06. Role: Biostatistician.

Newton TL (PI). Intimate partner violence and inflammatory processes. National Institute on Aging, National Institutes of Health: R21 AG0249002-02. Role: Biostatistical consultant.

Pfeiffer MP (PI). A Telehealth Intervention to Improve Symptom Management for Patients with Head and Neck Cancer. National Institutes of Health. RFA-CA-07-013. Role: Biostatistician.

Ruble LA (PI). Outcomes of Teacher Training on Autism. Department of Health and Human Services, Public Health Service. PAR-03-078. Role: Co- investigator.