RECURRENT EVENT STUDIES: EFFICIENT PANEL DESIGNS AND JOINT MODELING OF EVENTS AND SEVERITIES

by

Elizabeth Juarez-Colunga

B. Sc., Autonomous University of Queretaro, Mexico, 2002

M. Sc., National Autonomous University of Mexico, 2004

THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE DEPARTMENT

OF

STATISTICS AND ACTUARIAL SCIENCE

FACULTY OF SCIENCE

SIMON FRASER UNIVERSITY

Summer 2011

APPROVAL

Name: Elizabeth Juarez-Colunga

Degree: DOCTOR OF PHILOSOPHY

Title of Thesis: Recurrent Event Studies: Efficient Panel Designs and Joint Modeling of Events and Severities

Examining Committee: Dr. Tim Swartz (Chair)

Dr. Charmaine Dean
Senior Supervisor

Dr. Robert Balshaw
Supervisory Committee

Dr. Leilei Zeng
Supervisory Committee

Dr. Richard Lockhart
Internal External Examiner

Dr. Karen Kopciuk
Southern Alberta Cancer Research Institute
External Examiner

Date Approved: August 12, 2011

Declaration of
Partial Copyright Licence

Abstract

Recently there has been tremendous growth in the use and interest of longitudinal data, particularly because of the development of large scale investigations which are conducted to study different aspects of the dynamics of a population over time (for instance, the Canadian National Longitudinal Study of Children and Youth (Statistics Canada, 1996)). Recurrent event data are a type of longitudinal data which occur in many fields. Such data arise when an event repeats over time, and are common especially in medicine and reliability. Sometimes, in addition to the occurrence of the event, there is also information which reflects the severity of the event; this is called a mark. In this thesis, we develop efficient designs for longitudinal recurrent event studies, and we also develop methods to model recurrent events with marks.

In longitudinal recurrent event studies, sometimes partial information on the counting process, such as the number of events occurring in specific intervals, called panel data, provides nearly the same precision for estimation of treatment effects as full information based on data from continuous observation of the process. We compare the efficiency of the analysis of such panel data with respect to the analysis of data recorded as times of recurrences, and we articulate conditions for efficient panel designs where the focus is on estimation of a treatment effect when adjusting for other covariates. We model the recurrent intensity through the common proportional intensity framework, with the treatment effect modeled flexibly as piecewise constant over panels, or groups of panels. We provide some important considerations for the design of efficient panel studies.

The thesis also develops methods for situations where marks, denoting a measure of prognostic factors or severity of the event, are also recorded. Often, there is an association between the recurring processes of events and their marks. We model these outcomes jointly through the use of shared or linking random effects, and investigate biases resulting in analyses of the outcomes when they are not modeled jointly. This analysis of joint outcomes is motivated by a study of healthy menstruating women prior to hysterectomy/ovariectomy for benign disease.

Acknowledgments

I would like to express my deepest gratitude to my senior supervisor, Dr. Charmaine Dean for mentoring me in so many ways. She has greatly influenced my life, and I will be forever grateful. I would also like to thank the Department of Statistics and Actuarial Science for providing a great environment to pursue graduate studies. A special thanks to Dr. Richard Lockhart, Dr. Robert Balshaw, and Dr. Leilei Zeng for their support and advice. Thanks also to my fellow graduate students for their support, friendship and companionship; special mention goes to Carolyn Huston, Jean Shin, Kelly Burkett, Simon Bonner, Darby Thompson, Laurie Ainsworth, Cindy Feng, Ryan Lekivetz, Saman Muthukumarana, and Chunfang Lin.

I appreciate the support, encouragement, and friendship of Federico O'Reilly, Silvia Ruiz, Alberto Contreras, and Giovani Silva.

For their financial support, I would like to thank the National Council of Science and Technology of Mexico (CONACyT).

Finally, I would like to thank my family for their love and support. Shad, without you this would definitely had not been possible; you are my source of strength.

Contents

Approval ii

Abstract iii

Acknowledgments v

Contents vi

List of Tables viii

List of Figures x

1 Introduction 1

1.1 Recurrent Event Panel and Continuous Data 1

1.2 Poisson Process Models 2

1.3 Bayesian Methods 4

1.4 Outline 4

1.4.1 Efficient Designs for Recurrent Event Studies 4

1.4.2 Illustration of Efficient Designs and Power Considerations 5

1.4.3 Joint Analysis of Recurrent Events and Severities 5

1.4.4 Future Work 5

2 Efficient Designs for Recurrent Event Studies 6

2.1 Introduction 6

2.2 Asymptotic Relative Efficiency of Panel Studies 8

2.2.1 Efficiency Comparisons for Mixed Non-homogeneous Poisson Processes with Piecewise Constant Treatment Effects 8
2.3 Efficiency under imbalanced designs 18
2.3.1 Efficiency under imbalanced designs: 1-segment 19
2.3.2 Efficiency under imbalanced designs: 2-segments 28
2.4 Discussion 32
2.5 ARE of Panel Data Estimators 32
2.6 Notation 36

3 Illustration of Efficiency and Power 41
3.1 Introduction 41
3.2 Bladder Cancer Data 41
3.3 Measures of Balance 42
3.4 Analysis of Bladder Cancer Data 47
3.5 Simulation: Power Considerations 47
3.6 Discussion 50

4 Joint Analysis of Recurrent Events and Severities 52
4.1 Introduction 52
4.2 Hormone Therapy Study 54
4.3 Three-component Joint Model for Zero Heavy Counts and Severities 56
4.4 Analysis of Hormone Therapy Study 58
4.5 Investigation of Bias in Misspecified Shared Frailty Models 66
4.6 Summary 68

5 Future Work 71
5.1 Two-stage Designs for Recurrent Events 71
5.2 Assessment of the Use of Scoring Approaches for Analysis of Counts and their Severities 72
5.3 Multivariate Marked Recurrent Events with Exposure 73

Bibliography 75

List of Tables

2.1 Scenarios considered for illustration of AREs. The name of the scenario reflects the variable for which there is balance, and '*' denotes joint balance. 21

2.2 Scenarios considered for illustration of AREs. The name of the scenario reflects the variable for which there is balance; '*' denotes joint balance. A description of the scenarios is provided in Table 2.1. 22

3.1 Measures of balance in the bladder cancer data. The weighted differences of the w_i's, $l_{w,j}$, the weighted differences of z_i's, $l_{z,j}$, and the overall measure of balance l_j correspond to model-based measures; l_j^{**} and l_j^{*} correspond to the sum of squared differences of proportions of individuals per strata in the two treatment groups. For the 2-segments case, l_j^1 and l_j^2 correspond to measures of balance in segment 1 and 2, respectively; l_j corresponds to the measure of balance for the time-weighted overall mean treatment effect. Note that $r_{z,w}$ and the ratios H_p/H_f do not vary by treatment group, so they are displayed only once; 2p, 4p, and 8p denote 2, 4, and 8 panels, respectively. 46

3.2 Parameter estimates and their standard errors, resulting from the quasi-likelihood fit to the modified bladder cancer data when modeling the treatment effects as piecewise constant with 1 and 2 segments. 48

3.3 Parameter estimates and their standard errors, resulting from the quasi-likelihood fit to the original bladder cancer data when modeling the treatment effects as piecewise constant with 1 and 2 segments. 49

3.4 Design effects for different designs for a similar scenario as the bladder cancer dataset ($\alpha = 1; \beta_1 = -3.5; \beta_2 = -0.5; \gamma = 0; \tau = 0.9$) 50

4.1 Estimated linear predictors of logit(π_{it}) at 12, 24, 36, and 50 weeks from M_1 and M_3 for the hormone therapy analysis, and the simulation investigation. The panel labelled *HT Study* refers to the hormone therapy analysis, while that labelled *Simulated Data* refers to the investigation described in Section 4.5 64

4.2 Estimates of variance components and linking parameters in all models. The 95% credible interval limits are displayed under CI lower and CI upper. 65

4.3 Estimated correlations from joint models and model diagnostics. 66

4.4 Estimates of linear predictors of logit(π_{it}) at 12, 24, 36, and 50 weeks from the fits of separate analyses when the true model includes a shared frailty term with different degrees of correlation between each pair of random effects. Treatment is denoted as *Trt*. 69

List of Figures

2.1 Exploration of the ratio H_p/H_f for a Weibull intensity function ($\alpha t^{\alpha-1}$) with shape parameter $\alpha = 1.3$ (shown in a)). The ratio H_p/H_f as a function of the two panel followup times T_1 and T_2 ($T_2 > T_1$) is displayed in b). An efficient choice of T_1 and T_2 is suggested by the large values of H_p/H_f achieved near $T_1 = 6$ and $T_2 = 25$. 17

2.2 Illustration of panel designs for the treatment group in the scenario Bt*z (See Tables 2.2 and 2.1). Females are represented in the lighter colour (salmon), and males in the darker colour (blue). 20

2.3 Design scenarios for exploration of ARES. Females are represented in the lighter colour (salmon), and males in the darker colour (blue). Table 2.1 contains a detailed description of these scenarios, and Table 2.2 provides the definition of the design. 23

2.4 AREs of the estimate of the treatment effect, $\tilde{\beta_2}$, for different designs with 2, 4, or 8 panels, varying values of the covariate effect, γ (x-axis), and the overdispersion parameter, τ (y-axis). 25

2.5 AREs of the estimator of α, the shape parameter of the baseline intensity function, for different designs with 2, 4, and 8 panels, varying values of the covariate effect, γ (x-axis), and of the overdispersion parameter, τ (y-axis). . 26

2.6 AREs of the estimator of β_1, the overall mean baseline treatment effect, for different designs with 2, 4, and 8 panels, varying values of the covariate effect (x-axis), γ, and of the overdispersion parameter, τ (y-axis). 27

2.7 AREs of $\tilde{\beta}_2^1$ for the 2-segment model for different designs with 1 and 2 panels per segment. Parameter values are $\alpha = 1.1, \beta_1 = -3.5$ for the baseline intensity function; $\beta_2^1 = 0.2, 0.075, -0.050$ when the segment cut point is at 16, 24, and 32 months, respectively; $\beta_2^2 = -0.8, -0.925, -1.050$ when the segment cut point is at 16, 24, and 32 months, respectively. 29

2.8 AREs of $\tilde{\beta}_2^2$ for the 2-segment model for different designs with 1 and 2 panels per segment. Parameter values are $\alpha = 1.1, \beta_1 = -3.5$ for the baseline intensity function; $\beta_2^1 = 0.2, 0.075, -0.050$ when the segment cut point is at 16, 24, and 32 months, respectively; $\beta_2^2 = -0.8, -0.925, -1.050$ when the segment cut point is at 16, 24, and 32 months, respectively. 30

2.9 AREs of the estimate of the time-weighted mean treatment effect, $\tilde{\delta}_2$, based the 2-segment model for different designs with 1 and 2 panels per segment. Parameter values are $\alpha = 1.1, \beta_1 = -3.5$ for the baseline intensity function; $\beta_2^1 = 0.2, 0.075, -0.050$ when the segment cut point is at 16, 24, and 32 months, respectively; $\beta_2^2 = -0.8, -0.925, -1.050$ when the segment cut point is at 16, 24, and 32 months, respectively. 31

3.1 Bladder cancer data: the plot displays the termination times for the original data (a), and modified data (b). For the display of the covariates and individual-specific rates, the length of the lines are proportional to the size of the largest tumor, the number of tumors, and the rate of events, respectively; additionally, lighter colors are used to highlight large values of these variables. 44

3.2 Distributions of variables in the bladder cancer data: original data (a), and modified data (b). The smooth line corresponds to a smoother based on a Gaussian kernel. The mean and the median are displayed on the plots. . . . 45

3.3 Power of the test that $\beta_2 = 0$ for design Bz_b with parameters $\alpha = 1, \beta_1 = -3.5, \gamma = 0, \tau = 0$. The different curves denote the power for different sample sizes and designs; power with a sample size of 480*1.5 and 12000*1.03 are shown in a) and b), respectively. 51

4.1 Hot flush episodes showing the count level (upper graph) and the severity level per day (lower). A dark shade of red indicates larger counts and high severities. Time is given in days. 55

4.2 Estimated predictors from the fits of models M_1 (left panel) and M_3 (right panel). Rows display estimates of $\text{logit}(p_{it})$ (top), $\log(\mu_{it})$ (middle) and $\text{logit}(\pi_{it})$ (bottom). Estimates for the CEE group/ MPA group are displayed in green/brown solid lines with green/brown dashed lines indicating 95% confidence intervals. 60

4.3 Posterior means of the random effects for each individual obtained from the fit of M_3. Only records with events are shown. The numbers in the figures denote the bimester to which the random effect belongs; for instance, '1' refers to the first two months. 61

4.4 Estimated linear predictors at 12, 24, 36, and 50 weeks from M_1, M_2, and M_3. 63

4.5 a) Linear predictors of $\text{logit}(\pi_{it})$: black lines correspond to true values, and green (CEE) and brown (MPA) to the median of the estimated values from the fit of M_3 over the 1000 simulations. b) Linear predictors of $\text{logit}(\pi_{it})$: black lines correspond to true values, and different shades of gray correspond to means of estimates from the fit of the separate model over 250 simulations, where the pair-wise correlations between model components are $0, 0.10, 0.50$, and 0.90. 67

Chapter 1

Introduction

Recurrent event data arise in many fields where an event may repeat over time, perhaps especially in applications in medicine and reliability. The methods commonly used to model these are based on point processes. Many times, information on the process is only recorded at certain points in time, giving rise to what is termed here *panel data*. In this thesis we develop efficient designs for recurrent event studies giving rise to panel data, and we develop methods to model recurrent events where additional outcome data pertaining to the events are available in the form of *marks*. For example, marks may refer to the severity of the events.

There has been tremendous growth in the use of longitudinal panel studies for monitoring processes over time, and this has been accompanied by rapid developments in methodological tools for handling panel data. However, little attention has been devoted to the design aspects of such studies. A goal of this thesis is to fill that gap. In particular, we determine conditions for efficient estimation of a treatment effect from the analysis of panel data with respect to an analysis using continuous followup of individuals.

As well, we address methods for the joint analysis of counting processes and their marks and discuss issues related to such analyses.

1.1 Recurrent Event Panel and Continuous Data

With recurrent event panel studies each individual is observed at specific points in time and the number of occurrences of an event of interest between these followup times is recorded.

For individual i, let the total observation period be $[T_{i,o}, T_{i,e_i}]$ and the *panel followup times*, the times at which individuals are examined for recurrences, be $T_{i,o} < T_{i,1} < T_{i,2} \ldots < T_{i,e_i}$. Let M denote the sample size of the study, so i ranges from 1 to M. Panel p refers to the interval $(T_{i,p-1}, T_{i,p}]$, $p = 1, \ldots, e_i$, so e_i is the number of followup times for individual i. An alternative to panel data studies of recurrent events is continuous followup, where the time of occurrence of each event is recorded; we denote these times, in relation to the panels of followup mentioned above, as ω_{ipl}, where p denotes the panel in which the event occurred and l indexes the event within panel p, $p = 1, \ldots, e_i$, $l = 1, \ldots, n_{ip}$, n_{ip} being the number of events observed within panel p for individual i.

In the study of counting processes, an important consideration is an appropriate time scale. This is often calendar time, particularly in medical studies, but different scales may be appropriate in other applications; for example, the number cycles a particular machinery part is in use, or the mileage of a car. The time of origin is another important consideration; in medical studies this is often the beginning of treatment or the time of diagnosis. In this thesis, we will use calendar time with $T_{i,o} = 0$.

1.2 Poisson Process Models

Methodology for recurrent events can be found in several key references, including Ross (1996), Cox and Isham (1980), Andersen et al. (1993), and more recently Cook and Lawless (2007) and Aalen et al. (2008), with developments using a variety of approaches including Markov and semi-Markov approaches (Dabrowska et al., 1994), the analysis of gap times between events (Gail et al., 1980), and Poisson processes (Lawless, 1987). In this thesis we will focus on Poisson and related processes.

A Poisson process is a stochastic process with events occurring randomly over time. Let $N(t)$ be the number of events that occur in $[0,t]$. The corresponding counting process $\{N(t), t \geq 0\}$ records the cumulative number of events. More generally, let $N(s,t)$ denote the number of events over the interval $(s,t]$. A counting process is said to be a nonhomogeneous Poisson process with intensity $\lambda(t)$ if and only if:

1. $N(0) = 0$,
2. $Pr\{N(t) - N(t-h) = 1 | H(t-h)\} = \lambda(t)h + o(h)$

3. $Pr\{N(t)-N(t-h)>1|H(t-h)\}=o(h)$

for small h and $t>0$, and where $o(h)$ is such that $o(h)=\lim_{h\to 0}o(h)/h=0$. The history of the process, $H(t)$, denotes the record of all the events to t, $H(t)=\{N(u):0\leq u\leq t\}$. One of the key properties of the Poisson process is that the number of events in a window of time has a Poisson distribution, i.e.

$$Pr\{N(t)-N(t-h)=n\}=\frac{\left\{\int_{t-h}^{t}\lambda(u)du\right\}^{n}\exp\left\{-\int_{t-h}^{t}\lambda(u)du\right\}}{n!}. \tag{1.1}$$

Thus $N(t)$ has a Poisson distribution with mean $\Lambda(t)$, where

$$\Lambda(t)=\int_{0}^{t}\lambda(u)du \tag{1.2}$$

is termed the cumulative intensity function or the cumulative mean function. The likelihood based on n event times $0=t_0\leq t_1\leq t_2\leq\ldots\leq t_n\leq\tau$ occurring during the time $[0,\tau]$ for fixed τ is

$$\begin{aligned}L&=\left[\prod_{j=1}^{n}\lambda(t_j)\exp\left(-\int_{t_{j-1}}^{t_j}\lambda(u)du\right)\right]\exp\left(-\int_{t_n}^{\tau}\lambda(u)du\right)\\&=\left[\prod_{j=1}^{n}\lambda(t_j)\right]\exp\left(-\int_{0}^{\tau}\lambda(u)du\right).\end{aligned}$$

Methods for recurrent events are specified through the intensity function $\lambda(t)$, i.e. the instantaneous probability of recurrence. Another way to conceptualize recurrent event data is as a generalization of survival data (Lawless, 2003). Recall survival data models may be specified through the hazard function, i.e. the instantaneous probability of death. Models for hazard functions for survival data may often be usefully employed for the intensity function for Poisson processes.

Sometimes, in addition to the counting process there is a variable describing the event which reflects perhaps the severity of the event or some other outcome related to the event. This type of process is known as a *marked point process*, and is discussed in depth in Andersen et al. (1993) and Martinussen and Scheike (2006).

1.3 Bayesian Methods

Sometimes there is a need for accommodating complex correlation structures for several parameters; for instance, when observing two or more processes that are correlated, and each of the processes are also correlated over time. A computationally effective approach for such investigations is the use of Markov chain Monte Carlo techniques in a Bayesian framework of analysis. Bayesian methods combine prior belief with data information to obtain what is termed posterior distributions of the parameters; such posterior distributions are used to interpret effects and test hypotheses. There are several key references on Bayesian methodology; from a theoretical perspective (for example, Bernardo and Smith (1994)) to practical considerations (for example, Gelman and Rubin (2004); Carlin and Louis (1996)), as well as on computational approaches (for example, Ntzoufras (2009)).

1.4 Outline

This thesis contains three main research contributions as discussed in the following subsections.

1.4.1 Efficient Designs for Recurrent Event Studies

In longitudinal recurrent event studies, sometimes partial information on the counting process, such as the number of events occurring in specific intervals, i.e. panel data, provides nearly the same precision for estimation of treatment effects as full information based on data from continuous observation of the process. In Chapter 2, we compare the efficiency of the analysis of such panel data with respect to the analysis of data recorded as times of recurrences, and we articulate conditions for efficient panel designs where the focus is on estimation of a treatment effect when adjusting for other covariates. We model the intensity for the recurrent event process through the common proportional intensity framework, with the treatment effect modeled flexibly as piecewise constant over panels, or groups of panels. We provide some important considerations for the design of efficient panel studies.

1.4.2 Illustration of Efficient Designs and Power Considerations

In Chapter 3 we conduct an analysis of the recurrence of tumors in patients with bladder cancer (Andrews and Herzberg, 2000). We discuss the required sample size such that a panel design will achieve the same power in detecting a treatment effect as continuous followup. In addition, we describe the use of diagnostic measures which provide insight into the efficiency expected from a panel design.

1.4.3 Joint Analysis of Recurrent Events and Severities

Chapter 4 develops methods for situations where, in addition to the counts of events, we also record a mark, denoting a measure of prognostic factors or severity of the event. In many situations there is an association between the recurring processes and their marks. This occurs in the motivating example for this work, a study of the effects of two treatments in reducing vasomotor symptoms in a hormone therapy study including healthy menstruating women prior to hysterectomy/ovariectomy for benign disease. We model the event counts and their severities jointly through the use of shared or linking random effects, and investigate biases resulting in analyses where the outcomes are not modeled jointly.

1.4.4 Future Work

This thesis closes with a discussion of future work emerging from extensions of the methods considered as well as new topics of interest in the study of recurrent events.

Chapter 2

Efficient Designs for Longitudinal Recurrent Event Studies with Missing Data

2.1 Introduction

Considerable attention has been devoted to the analysis of recurrent event data, that is, data arising from counting processes governing the repeated occurrence of events over time. Much of this literature assumes continuous followup of individuals (Cook and Lawless, 2007), with developments related to Poisson processes (Lawless, 1987), the analysis of gap times between events (Gail et al., 1980), and a variety of Markov and semi-Markov approaches (Dabrowska et al., 1994). Additionally, there has been a rise in interest in panel data, that is, data arising as counts of events over time intervals, or panels. Panel data are common in clinical studies where for ethical or practical reasons only the number of episodes in an interval of time are recorded, as, for example, in the study of treatments for epilepsy (Thall and Vail, 1990). They are also conveniently adopted for national longitudinal surveys such as the Canadian National Longitudinal Study of Children and Youth (Statistics Canada, 1996), and the U.S. National Longitudinal Surveys on Labour Statistics (U.S. Department of Labor, Bureau of Labor Statistics, 2005). Sun et al. (2009) remark that panel data have considerable potential for understanding the underpinnings of complex

models.

Due to the widespread availability of longitudinal data, there has been a corresponding increase in the development of methods for handling such data; however, there are few studies emphasizing design effects and optimal design strategies, especially relating to the use of panel data in lieu of data from continuous followup. Such strategies are fundamental to better utilization of funds, especially in times of fiscal constraints, when national agencies implementing long-term longitudinal studies are required to justify their merit. Dean and Balshaw (1997) studied the efficiency of the analysis of counts versus the analysis of event times, as would occur through continuous followup. Their focus was on the analysis of aggregated counts over the full followup period in a simple k-sample study with a fixed treatment effect. Here we extend this substantially to examine how design effects influence efficiency of the analysis of panel event counts relative to the analysis of actual event times. We derive conditions which yield high efficiency of the estimator of a treatment effect in a k-sample study, while adjusting for other covariates. We permit the treatment effect to vary over panels, or groups of panels, and adopt a proportional intensity model where the treatment effect is piecewise constant over panel segments. Our principal aim is identification of conditions for full efficiency of the treatment effect but we also consider efficiencies of other model parameter estimates. Additionally, we illustrate efficiency losses for a variety of designs to demonstrate how sharply efficiencies change with different design conditions.

In Section 2.2, we derive the asymptotic relative efficiency of estimates of treatment effects. We provide efficiencies for several designs in Section 2.3, and provide, in Section 2.4, a summary and discussion of the methods presented. Section 2.5 contains a proof of the main theorem used in the discussions in Sections 2.2 and 2.3. *For convenience, we note here that a list of the notation used in this chapter is assembled in Section 2.6 (see page 36).* We have included these latter two sections at the end of the chapter so as to preserve the flow of the main ideas and concepts through the earlier sections.

2.2 Asymptotic Relative Efficiency of Panel Studies

2.2.1 Efficiency Comparisons for Mixed Non-homogeneous Poisson Processes with Piecewise Constant Treatment Effects

The proportional intensity model is widely used in part because of its tractability (Andersen and Gill, 1982). Sometimes, however, the need for a more flexible form of incorporation of a covariate effect arises, for instance, when treatment effects change over time (Sun et al., 2008). This might be the case when the benefit of the treatment is thought to be delayed until a certain period of exposure has been achieved. We consider a generalization of the usual proportional intensity model with a fixed treatment effect that allows the treatment effect to vary over groups of panels, so the treatment effect is piece-wise constant over specific time segments. We assume these segments are predefined, and each of them contains at least one panel.

Consider a comparison of k treatments, where m_j individuals are given treatment j, and let $M = \sum_{j=1}^{k} m_j$, the total sample size of the study. Let $\{N_i(t), t \geq 0\}$ denote the counting process monitoring the occurrence of events for subject i, $i = 1, \ldots, M$. The total observation period for all individuals is divided into S segments, defined by $T^1, T^2, \ldots, T^S$, over which the treatment effects are expected to be piecewise constant. Segment s is $(T^{s-1}, T^s]$, $s = 1, \ldots, S$; these might indicate periods over which changes in efficacy or response to the treatment are expected. As such, they are defined by important panel followup times at which the design requires all individuals to be seen, unless they have been lost to followup or censored. Individuals are expected to be seen more frequently than only at these times, following some suggested planned pattern of followups; these might be individual-specific. Data may also be aggregated over two or more consecutive panels; in this case we simply redefine the planned followup times to match what actually occurred. Missing panel counts are also possible. All these situations are accommodated by the likelihood development herein.

Each individual is observed up to time T_{e_i} referred to as the *termination time* for individual i. Let the observation process $\{Y_i(t), t \geq 0\}$ be 1 if individual i is under study at t and 0 otherwise, and assume that $\{Y_i(t), t \geq 0\}$ is independent of the counting process $N_i(t)$. Thus the observed counting process may be defined as $\bar{N}_i(t) = \int_0^t Y_i(u) dN_i(u)$, $t \geq 0$, $t \in (0, T_{e_i}]$

for individual i, $i = 1,\ldots,M$. Individual-specific *panel followup times* in segment s are denoted $(T^s_{i,1}, T^s_{i,2}, \ldots, T^s_{i,e^s_i})$, where, if individual i has not been lost to followup, $T^s_{i,e^s_i} = T^s$. Panel counts in segment s for individual i are denoted as $n^s_{ip} = \bar{N}_i(T^s_{i,p}) - \bar{N}_i(T^s_{i,p-1})$, $p = 1,2\ldots,e^s_i, s = 1,2,\ldots,S$, with the total aggregated count for individual i in segment s being $n^s_{i+} = \sum_{p=1}^{e^s_i} n^s_{ip}$.

We focus here on creating efficient study designs for estimation of a treatment effect. We define the covariate $\mathbf{x}_i$ as a $k \times 1$ treatment indicator vector for the i-th individual, such that $x_{i1} = 1$ represents an intercept term, and $x_{ij} = 1$ if individual i received treatment j, or 0 otherwise, $j = 2,\ldots,k$. The $d_z \times 1$ vector $\mathbf{z}_i$ contains remaining covariates which are to be adjusted for in the analysis.

The counting process $N_i(t)$ is modeled as a Poisson process with intensity function

$$\lambda_i(t) = \nu_i \rho(t;\alpha) \exp\left\{ \mathbf{x}_i^{'} \left[\sum_{s=1}^{S} \beta^s I_{(T^{s-1},T^s]}(t) \right] + \mathbf{z}_i^{'}\gamma \right\}, \tag{2.1}$$

given ν_i, an individual-specific random effect accounting for the common phenomenon of overdispersion; the function $I_{(T^{s-1},T^s]}(t)$ is an indicator function for $t \in (T^{s-1}, T^s]$; and ρ is a twice differentiable baseline intensity function depending on the parameter α with dimension d_α. The parameters $\beta^s = (\beta^s_1, \beta^s_2, \ldots, \beta^s_k)'$ and γ correspond to the treatment effects in segment s, and the parameters governing the effects of the covariates to be adjusted for in the analysis, $s = 1,\ldots,S$, respectively. We have parameterized the β's so that treatment effects are measured relative to treatment 1. Hence, β^s_1 reflects the overall baseline level of the intensity function $\lambda_i(t)$ in segment s, when the z_i's are centered, whereas α measures the shape of the intensity function $\rho(t,\alpha)$. Additionally, we may take $\mathrm{E}(\nu_i) = 1$ without loss of generality; let $\mathrm{var}(\nu_i) = \tau$. If we denote the mean of the total aggregated counts to be $\mu^s_{i+} = \mathrm{E}(n^s_{i+})$ in segment s, then its variance is of the form $\mu^s_{i+}(1 + \tau\mu^s_{i+})$. Writing the cumulative baseline intensity function in each segment s as $R^s_i = \int_{T^{s-1}}^{T^s} Y_i(t)\rho(t;\alpha)dt$, then $\mu^s_{i+} = R^s_i \exp(\mathbf{x}_i^{'}\beta^s + \mathbf{z}_i'\gamma)$. Similarly, defining the cumulative baseline intensity function in panel period p as $R^s_{ip} = \int_{T^s_{i,p-1}}^{T^s_{i,p}} Y_i(t)\rho(t;\alpha)dt$, we have $\mu^s_{ip} = \mathrm{E}(n^s_{ip}) = R^s_{ip}\exp(\mathbf{x}_i^{'}\beta^s + \mathbf{z}_i'\gamma)$.

Our objective is to investigate the efficiency of panel studies with respect to continuous followup, focusing on the estimation of the treatment effect β while adjusting for other covariates. A popular inferential technique, and the one used in this chapter, is quasi-likelihood analysis, where the mean and variance specification is sufficient to formulate

estimating equations for the parameters, and also to obtain the asymptotic variance of the estimators of β, γ and α.

To motivate the use of specific estimating equations, which reduce to common forms in the absence of overdispersion, we describe the likelihood for both panel and continuous data. The likelihood over all the segments, based on either the full data or the panel data, has the form $\prod_{s=1}^{S} L^s$ where L^s corresponds to the likelihood in segment s, $s = 1, \ldots, S$. In the discussion below, for the purpose of simplicity in the derivation of the theorem regarding efficiency, we focus on one segment, and drop the superscript s; however, we provide the theorem in its more general form based on several segments.

Let $\theta = (\beta', \gamma', \alpha', \tau)'$, and let ω_{ipl} be the time of the l-th event, from the start of the study, for the i-th individual in panel period p, $i = 1, \ldots, M$, $p = 1, \ldots, e_i$, $l = 1, \ldots, n_{ip}$. The likelihood based on either the full data (subscripted by $d = f$) or the panel data (subscripted by $d = p$) factorizes as:

$$L_d(\theta) = L_{\alpha,d}(\alpha) L(\theta), \qquad d \in \{f, p\} \tag{2.2}$$

where

$$L_{\alpha,f}(\alpha) = \prod_{i=1}^{M} \prod_{p=1}^{e_i} \prod_{l=1}^{n_{ip}} \frac{\rho(\omega_{ipl}; \alpha)}{R_i}, \tag{2.3}$$

and

$$L_{\alpha,p}(\alpha) = \prod_{i=1}^{M} \left[\binom{n_{i+}}{n_{i1}, \ldots, n_{ie_i}} \prod_{p=1}^{e_i} \left(\frac{R_{ip}}{R_i} \right)^{n_{ip}} \right]; \tag{2.4}$$

$$L(\theta) = \prod_{i=1}^{M} \int_0^\infty (\nu_i \mu_{i+})^{n_{i+}} e^{-\nu_i \mu_{i+}} (n_{i+}!)^{-1} p(\nu_i) d\nu_i \tag{2.5}$$

is the likelihood for a mixed Poisson model based on the total counts observed for individual i. For example, if ν_i is gamma distributed, $L(\theta)$ becomes the negative binomial likelihood. Notice that if there is a single panel, $L_p(\theta)$ (see 2.2) will reduce to a simple mixed Poisson kernel, $L(\theta)$.

The factorization of $\lambda_i(t)$ into two components, one a function of α, the other a function of β, γ, α and τ is the key to the factorization of the likelihood in (2.2). We exploit this in deriving estimators, and in computing the asymptotic efficiency of the estimator of β, (and α) derived from L_p, with respect to the corresponding estimators derived from L_f.

The procedure we suggest for robust estimation is a type of quasi-likelihood (Wedderburn, 1974), which consists of replacing the contribution from $L(\theta)$ in the estimating equations by suitable estimating functions. Let g_d denote the full set of estimating equations for the panel ($d = p$) or the full ($d = f$) data. The estimating equation for the covariate effect $\eta_1 = (\beta', \gamma')'$ for the panel or the full data is

$$g_{\eta_1} = V' U_1 U_o^{-1} (\mathbf{n} - \mu) = \mathbf{0} \tag{2.6}$$

where $U_1 = \text{diag}\{\mu_{i+}, i = 1, \ldots, M\}$, $U_o = \text{diag}\{\mu_{i+}(1 + \tau\mu_{i+}), i = 1, \ldots, M\}$, $\mathbf{n} = (n_{1+}, \ldots, n_{M+})'$ is a vector of counts, and $\mu = (\mu_{1+}, \ldots, \mu_{M+})'$ is the vector of their expected values; $V = (X \;\; Z)$ combines the treatment indicators in X with the covariates in Z. Equation (2.6) arises from the usual quasi-likelihood function $(\partial\mu/\partial\eta_1)'\text{var}(\mathbf{n})^{-\mathbf{1}}(\mathbf{n} - \mu)$.

We obtain an estimating equation for α by combining $\partial \log L_{\alpha,d}/\partial\alpha$, $d = f, p$, with quasi-likelihood estimation, yielding

$$g_{\alpha,d} = \frac{\partial \log L_{\alpha,d}(\alpha)}{\partial\alpha} + W' U_1 U_o^{-1} (\mathbf{n} - \mu) = \mathbf{0}, \tag{2.7}$$

where W is a matrix with entries

$$w_{ia} = \frac{\partial \log R_i}{\partial \alpha_a}, \qquad i = 1, \ldots, M, \text{ and } a = 1, \ldots, d_\alpha. \tag{2.8}$$

There are several choices for the estimating equation of the overdispersion parameter τ, and our results concerning efficiency and design remarks in this chapter are not tied to the use of any specific estimator for τ. In our examples, we use the pseudo-likelihood estimator, which has been popular since its introduction by Davidian and Carroll (1987). It has performed well in simulation studies and has documented optimality properties (Nelder and Lee, 1992). The pseudo-likelihood estimating equation for τ is

$$g_\tau = \sum_{i=1}^{M} \frac{(n_{i+} - \mu_{i+})^2 - (1 - h_i)\mu_{i+}(1 + \tau\mu_{i+})}{(1 + \tau\mu_{i+})^2} = 0, \tag{2.9}$$

where $h_i = \text{diag}(U_1^{1/2} V_1' (V_1' U_1 V_1)^{-1} V_1' U_1^{1/2})$, $V_1 = (\;X \quad Z \quad W\;)$; h_i is the diagonal of the hat matrix and represents a correction to reduce small sample bias in this simple second moment equation.

Hence, when we have full data we solve $g_f = (g'_{\eta_1}; g'_{\alpha,f}, g_\tau)' = \mathbf{0}$ to yield an estimator of θ denoted by $\hat{\theta}$, and, when we use a panel design, the estimating equation becomes

$g_p = (g'_{\eta_1}; g'_{\alpha,p}, g_\tau)' = \mathbf{0}$ yielding the estimator $\tilde{\theta}$. In the following, estimators obtained from g_f are notated with a hat (e.g. $\hat{\alpha}$) while those obtained from g_p are notated with a tilde ($\tilde{\alpha}$).

For $\eta = (\beta', \gamma', \alpha')'$, we find that the information matrix concerning these parameters has the partitioned form (2.10) for $d = f$ and $d = p$, respectively:

$$I_d = \begin{pmatrix} V'UV & V'UW \\ W'UV & W'UW + H_d \end{pmatrix}, \tag{2.10}$$

with

$$U = U_1 U_o^{-1} U_1 = \text{diag}\left\{ u_i = \frac{\mu_{i+}}{1 + \tau\mu_{i+}}, i = 1, \ldots, M \right\}, \tag{2.11}$$

and where the terms H_f and H_p are segment-specific functions, and are defined in (2.12) and (2.13) when we return to a consideration of multiple segments.

Theorem 1 provides the asymptotic relative efficiency of the panel treatment estimates in each segment, $\tilde{\beta}^s$, and $\tilde{\alpha}$, relative to the estimators from the full data, $\hat{\beta}^s$ and $\hat{\alpha}$. It is sometimes of interest to evaluate the overall effect of a treatment over the study period, and Theorem 1 also provides the asymptotic relative efficiency of the estimate of the time-weighted mean effect of treatment j, $\tilde{\delta}_j = \sum_{s=1}^{S} \Delta^s \tilde{\beta}^s_j$, where $\Delta^s = (T^{s-1} - T^s)/T^S$. We first present the theorem and subsequently discuss interpretation in great depth. The theorem is notationally cumbersome; we focus on structures which represent the efficiencies so that the main elements stand out, and for clarity, we present results for scalar α and a single covariate Z. We provide, in Section 2.5, a proof of the case where α is an arbitrary vector, and covariate Z has a general structure, for a one-segment study. The derivation of the theorem for a study utilizing multiple segments follows in a straightforward manner. Additionally, to simplify the discussions herein, we adopt the following two conventions: (1) Because α is scalar, we denote the value w^s_{i1} (see 2.8) simply as w^s_i. (2) We let G^s_j be the set indexing individuals in treatment group j observed in segment s, $j = 1, \ldots, k$, $s = 1, \ldots, S$, and let $[u]^s_{j+} = \sum_{i \in G^s_j} u^s_i$ (see 2.11), $[uw]^s_{j+} = \sum_{i \in G^s_j} u^s_i w^s_i$, $[uz]^s_{j+} = \sum_{i \in G^s_j} u^s_i z_i$. Similarly, the total number of events observed for all individuals receiving treatment j in segment s, $\sum_{i \in G^s_j} n^s_{i+}$, is represented by $[n]^s_{j+}$.

The expressions for H_f and H_p in (2.10) are

$$H_f = \sum_{s=1}^{S}\sum_{i=1}^{M}\sum_{p=1}^{e_i^s} E\left\{\sum_{l=1}^{n_{ip}^s} -\frac{\partial^2 \log[\rho(\omega_{ipl}^s;\alpha)/R_i^s]}{\partial\alpha\partial\alpha'}\right\}; \tag{2.12}$$

$$H_p = \sum_{s=1}^{S}\sum_{i=1}^{M} E\left\{\sum_{p=1}^{e_i^s} -n_{ip}^s \frac{\partial^2 \log[R_{ip}^s/R_i^s]}{\partial\alpha\partial\alpha'}\right\}. \tag{2.13}$$

Theorem 1

a) The asymptotic relative efficiency (ARE) of $\tilde{\alpha}$ relative to $\hat{\alpha}$ is

$$\text{ARE}(\tilde{\alpha}) = 1 - \left(1 - \frac{H_p}{H_f}\right)\left(1 + \frac{E}{H_f}\right)^{-1}, \tag{2.14}$$

where

$$E = \phi_{w,w}\left(1 - \frac{\phi_{z,w}^2}{\phi_{z,z}\phi_{w,w}}\right), \tag{2.15}$$

$\phi_{z,w} = \sum_{s=1}^{S}\sum_{j=1}^{k}\sum_{i\in G_j^s} u_i^s \left(z_i - [uz]_{j+}^s/[u]_{j+}^s\right)\left(w_i^s - [uw]_{j+}^s/[u]_{j+}^s\right)$; and $\phi_{z,z}$ and $\phi_{w,w}$ are correspondingly defined. These quantities are sums of segment- and group-specific weighted variation and covariation of z_i and w_i^s, with weights u_i^s depending on the size of the mean term $\mu_i^s, s = 1,\ldots,S$.

b) The asymptotic relative efficiency of $\tilde{\beta}_j^s$ relative to $\hat{\beta}_j^s$ is

$$\text{ARE}(\tilde{\beta}_j^s) = 1 - \left\{\frac{(l_j^s)^2}{\left(l_{o,j}^s + \frac{(l_{z,j}^s)^2}{\phi_{z,z}}\right)(E + H_p) + (l_j^s)^2}\right\}\left(1 - \frac{H_p}{H_f}\right)\left(1 + \frac{E}{H_f}\right)^{-1}, \tag{2.16}$$

$$l_j^s = \sqrt{\phi_{w,w}}\left\{\frac{l_{w,j}^s}{\sqrt{\phi_{w,w}}} - \frac{l_{z,j}^s}{\sqrt{\phi_{z,z}}} r_{z,w}\right\}, \quad j = 1,\ldots,k, \tag{2.17}$$

where, for the baseline group, $j = 1$, $l_{o,1}^s = 1/[u]_{1+}^s$, $l_{a,1}^s = [ua]_{1+}^s/[u]_{1+}^s$, $a = w,z$. For the treatment groups, $j = 2,\ldots,k$, $l_{o,j}^s = 1/[u]_{1+}^s + 1/[u]_{j+}^s$, $l_{a,j}^s = [ua]_{j+}^s/[u]_{j+}^s - [ua]_{1+}^s/[u]_{1+}^s$, $a = w,z$; $r_{z,w} = \phi_{z,w}/\sqrt{\phi_{w,w}\phi_{z,z}}$.

c) The asymptotic relative efficiency of $\tilde{\delta}_j$ relative to $\hat{\delta}_j$ is

$$\text{ARE}(\tilde{\delta}_j) = 1 - \left\{ \frac{l_j^2}{\left(\sum_{s=1}^{S} (\Delta^s)^2 l_{o,j}^s + \frac{\left(\sum_{s=1}^{S} \Delta^s l_{z,j}^s\right)^2}{\phi_{z,z}} \right)(E + H_p) + l_j^2} \right\} \left(1 - \frac{H_p}{H_f}\right)\left(1 + \frac{E}{H_f}\right)^{-1}, \tag{2.18}$$

where $l_j = \sum_{s=1}^{S} \Delta^s l_j^s$, $j = 1, \ldots, k$.

2.1.1 Insight into Conditions for High Efficiency

Consider first the special case of a single segment, S=1, and no overdispersion, $\tau = 0$. The parameter estimates of β from the full data likelihood, L_f (see 2.2), satisfy

$$e^{\hat{\beta}_j} = \frac{[n]_{j+}}{e^{\hat{\beta}_1}[R(T_e;\hat{\alpha})e^{\mathbf{z}'\hat{\gamma}}]_{j+}} = \frac{\sum_{i \in G_j} n_{i+}}{e^{\hat{\beta}_1} \sum_{i \in G_j} R_i(T_{e_i};\hat{\alpha})e^{\mathbf{z}_i'\hat{\gamma}}}, \tag{2.19}$$

$j = 2, \ldots, k$, and $\exp(\hat{\beta}_1) = [n]_{1+}/[R(T_e;\hat{\alpha})\exp(\mathbf{z}'\hat{\gamma})]_{1+}$. Similarly, the estimates based on the panel data likelihood satisfy $\exp(\tilde{\beta}_1) = [n]_{1+}/\ [R(T_e;\tilde{\alpha})\exp(\mathbf{z}'\tilde{\gamma})]_{1+}$ and $\exp(\tilde{\beta}_j) = [n]_{j+}/(\exp(\tilde{\beta}_1)[R(T_e;\tilde{\alpha})\exp(\mathbf{z}'\tilde{\gamma})]_{j+})$, $j = 2, \ldots, k$. An important simple result is that if the set of termination times T_e's for each covariate stratum for individuals on treatment 1 are identical to those on treatment j, then the estimates of β_j from the full and panel analyses are identical,

$$e^{\hat{\beta}_j} = e^{\tilde{\beta}_j} = \frac{[n]_{j+}}{[n]_{1+}}, \quad j = 2, \ldots, k. \tag{2.20}$$

In this situation, the asymptotic variances of $\hat{\beta}_j$ and $\tilde{\beta}_j$ would also be identical, so analysis of the panel counts when the design uses any number of panels is fully efficient for estimation of β_j, $j = 2, \ldots, k$. Such a design is illustrated in Figure 2.3 a) (and discussed in Section 2.3.1; see also Tables 2.1 and 2.2 in that section for more details). There are two strata in each of the baseline and treatment groups, representing males and females, for illustrative purposes. Adjusting for gender in this analysis, the estimate of the treatment effect from this panel design is fully efficient. Even though there are more male dropouts earlier in the study, this does not affect the efficiency of the treatment effect; we will illustrate later that other types of imbalance will affect this efficiency.

More generally, calculation of (2.16) for a proposed design using a range of values for true treatment effects would, of course, give an indication of efficiencies. Here we describe conditions yielding fully efficient estimators, i.e. an ARE of unity. An important point throughout this discussion is that the form of the dependence of the counting process on α is left arbitrary, so these results hold for semiparametric models of the form (2.1).

1. Joint balance. If $l_j^s = 0$ (see 2.17), then the panel estimator $\tilde{\beta}_j^s$, $j = 2,\ldots,k, s = 1,\ldots,S$ is fully efficient. We refer to this condition as *joint balance* between the w_i^s's (or termination times T_{e_i}'s) and the covariates z. Consider the simple example illustrated in Figure 2.3 a), where the design has equal sample sizes, there is a single covariate, gender, and no overdispersion, $\tau = 0$. The gender-specific set of termination times are identical for the baseline and treatment groups. When $\tau = 0$ and there is no missing data, the simplest case in which *joint balance* is achieved for all segments occurs when termination times for each value of the covariate are identical for the baseline and treatment groups. It is also achieved, when $\tau = 0$, when there are r times as many individuals in the treatment group as in the baseline, and the termination times for each value of the covariate is an r-replicate in the treatment group of the corresponding values for the baseline group. In Figure 2.3 a), any panel design (for example, with panel followup times at 16, 32, 48, and 64 months and with two segments over 0-32 months and 32-64 months) will lead to a fully efficient estimator of the treatment effects β_2^s, in all segments s, after adjusting for gender. When there are missing data, *joint balance* requires that the observation period for individuals of each gender (or, more generally, for each value of the covariate vector) be identical (or r-replicates) in the treatment group as for the baseline group.

As suggested above, we use l_j^s, $s = 1,\ldots,S$ as our principal *measure of balance* for estimation of treatment effects; additionally, we partition l_j^s into components $l_{w,j}^s$, which reflects imbalance in termination times, $l_{z,j}^s$, which reflects imbalance in covariates, and $r_{z,w}$ (see 2.17), which modulates the effect of imbalance in the covariates on the ARE.

2. Balance. A weaker form of *balance* refers to having the same relative frequency of the termination times in each of the treatment and the baseline groups, without regard for how these are allocated over the covariate; this is denoted *balance in the termination times*. We will also consider the usual concept of *balance in the covariates*, where there are the same number of individuals in each covariate stratum, for example, the same

number of males and females, in this case without regard to any differences in how the termination times by treatment are distributed over gender. We remark here immediately that balance in the termination times is more important than balance in the covariates for achieving a high ARE of the estimate of the treatment effect; this is because of the modulating effect of $r_{z,w}$ in (2.17).

3. Overdispersion. With overdispersion, $\tau > 0$, the panel estimator of the treatment effect is fully efficient only when $\beta_j^s = 0$. However, we will illustrate later, that as long as *joint balance* holds, (in fact, even quite approximately) the estimator of β_j^s retains very high efficiency even if τ and β_j^s are far from zero.

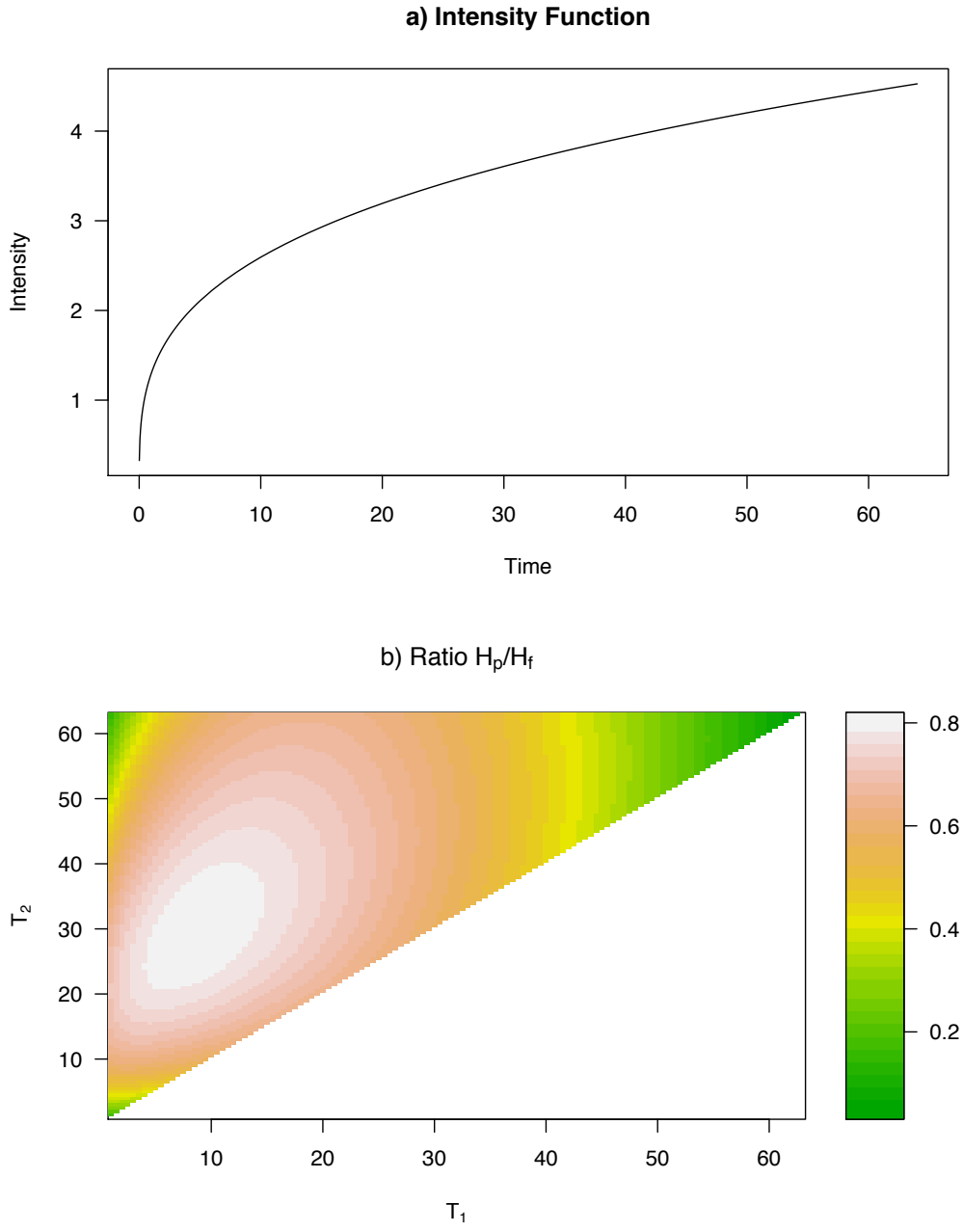
4. Effect of H_p. The AREs also depend on the ratio H_p/H_f, and this component plays a key role when the design is not balanced. The ratio increases rather quickly to 1 as the number of panels increases. We will illustrate this for some examples in Section 2.3. Another key remark is that the placement of the panel followup times affects the ARE through H_p; note the influential role of H_p/H_f in determining the efficiency of the shape parameter of the baseline function (see 2.14). Examination of this term can suggest optimal placement of the panel followup times. For example, considering 1 segment (S=1) and with 3 panels, for the Weibull intensity, we have $H_p = \sum_{i=1}^{M}\sum_{p=1}^{e_i}\mu_{ip}T_{i,p}^{\alpha}T_{i,p-1}^{\alpha}(\log T_{i,p} - \log T_{i,p-1})^2/(T_{i,p}^{\alpha} - T_{i,p-1}^{\alpha})^2$ and $H_f = \sum_{i=1}^{M}\sum_{p=1}^{e_i}\mu_{ip}/\alpha^2$. Figure 2.1 illustrates the ratio H_p/H_f for a 3-panel design for varying values of T_1 and T_2; the Weibull baseline is also displayed in this figure. An optimal choice of followup times, yielding large values of H_p/H_f is achieved here, for instance, for $T_1 = 6$ and $T_2 = 25$. Choosing values of T_1 which are quite large, well beyond the early sharp changes in the intensity, for example, greater than 30, yields lower efficiencies no matter the choice of $T_2 > T_1$.

For some common parametric forms of the baseline intensity $\rho(t)$, H_f can be written as

$$H_f = \sum_{s=1}^{S}\sum_{i=1}^{M}\sum_{p=1}^{e_i}\mu_{ip}^s\psi_f^s(T_{i,p-1}^s, T_{i,p}^s;\alpha) \tag{2.21}$$

where $\psi_f^s(T_{i,p-1}^s, T_{i,p}^s;\alpha)$ is a function of the $T_{i,p}^s$'s and α only; H_p also has the same structure with ψ_f^s on the right hand side replaced with ψ_p^s. This form for H_f is possible, for example, when $\rho(t)$ is Weibull (αt^{α}) or exponential ($\exp(\alpha t)$). When (2.21) holds, further observations regarding the AREs can be made as discussed in the following two remarks.

Figure 2.1: Exploration of the ratio H_p/H_f for a Weibull intensity function ($\alpha t^{\alpha-1}$) with shape parameter $\alpha = 1.3$ (shown in a)). The ratio H_p/H_f as a function of the two panel followup times T_1 and T_2 ($T_2 > T_1$) is displayed in b). An efficient choice of T_1 and T_2 is suggested by the large values of H_p/H_f achieved near $T_1 = 6$ and $T_2 = 25$.



5. Size of the treatment effects. When H_f has the form (2.21) and $\tau = 0$, the ARE of $\tilde{\beta}_j^s$ becomes large as $|\beta_j^s|$ increases, $j = 2, \ldots, k$. If there is a single segment, $S = 1$, the ARE of $\tilde{\beta}_j^s$ does not depend on β_1^s.

6. Baseline intensity estimators. The ARE of $\tilde{\alpha}$ will be high when H_p/H_f is large, and ARE($\tilde{\beta}_1^s$) will be high when the difference in the weighted means of the w_i^s's and z_i's is small. If (2.21) holds, $\tau = 0$ and $S = 1$, these AREs do not depend on β_1^s, the overall level of the mean count. Under conditions for *joint balance*, the ARE of $\tilde{\beta}_1^s$ increases with increasing values of $\sum_{j=2}^{k} \exp(\beta_j^s)$ and the ARE of $\tilde{\alpha}$ does not depend on β^s.

7. Effect of E. The ARE of $\tilde{\alpha}$ (2.14) will also be large when E (2.15) is large. The term E denotes the weighted variation of w_i^s's rescaled by a measure of the association between the w_i^s's (a function of the termination times) and z_i's. If the weighted variation of w_i^s's is large or the measure of association is small then E becomes large and so too the ARE of $\tilde{\alpha}$. Intuitively, when event times are not available and panel counts are used, the more variation there exists in the termination times, the more information there will be regarding the shape of the baseline intensity (parametrized by α).

8. Time-weighted mean. When there is *joint balance*, the estimator of the time-weighted mean, $\tilde{\delta}_j$, will be fully efficient. Intuitively, under imbalanced scenarios, imbalance in one segment may be compensated by different imbalances in another segment, so that the overall estimator may become less affected by imbalance.

2.3 Efficiency under imbalanced designs

In Subsections 2.3.1 and 2.3.1, we consider imbalanced designs to ascertain how sensitive efficiencies may be to departures from the condition of *joint balance* when considering 1 and 2 segments.

Consider a study with two treatment groups, and a single covariate, gender, with 2, 4 and 8 equally spaced scheduled panel followup times (Figure 2.2) over a period of 64 months. Imbalance in the data could be interpreted conceptually as individuals abandoning the study before the 64 months, as displayed in Figure 2.2, or, by staggered entry of individuals; at the scheduled panel followup times and at the end of followup, we record information on events observed between followup visits. We assume a Weibull hazard for

the baseline intensity function.

In this section we numerically study the efficiency of estimators of treatment effects under imbalanced designs, i.e. designs for which joint balance does not hold.

2.3.1 Efficiency under imbalanced designs: 1-segment

Although the primary interest here is the efficiency of the estimator of the treatment effect, we also comment on the efficiency of estimates of α and β_1, which describe the shape and the overall level of the intensity function.

Table 2.1 below, and Table 2.2 list six contrasting designs (labelled Bt*z, Btz, Bt, Bz_a, Bz_b, B), where the nomenclature reflects whether balance is achieved over the treatment and control groups in the covariates (z) or the termination times T_{e_i}'s (t). Figure 2.3 illustrates all these designs. The event times displayed by $+$ signs in Figure 2.3 are one realization of the random generating process. Table 2.1 provides a description of the designs, while Table 2.2 provides details on the numbers of individuals observed and their panel followup times. In Table 2.2, two designs achieve *balance in the covariates* (Bz) with different levels of imbalance in the termination times: these are denoted Bz_a and Bz_b. All designs spread the dropouts equally at $16, 24, 32, 40, 48$, and 56 months. For example, for the design with *joint balance*, Bt*z, where 10% of females and 60% of males drop out in each of the control and treatment groups, the loss to followup occurs as one female and six male dropouts at each of $16, 24, 32, 40, 48$ and 56 months in each of these groups. For clarity, Figure 2.2 illustrates the loss to followup for the control or treatment groups and the $2, 4$, and 8 equally spaced panel followup times for the design Bt*z.

Figure 2.2: Illustration of panel designs for the treatment group in the scenario Bt*z (See Tables 2.2 and 2.1). Females are represented in the lighter colour (salmon), and males in the darker colour (blue).

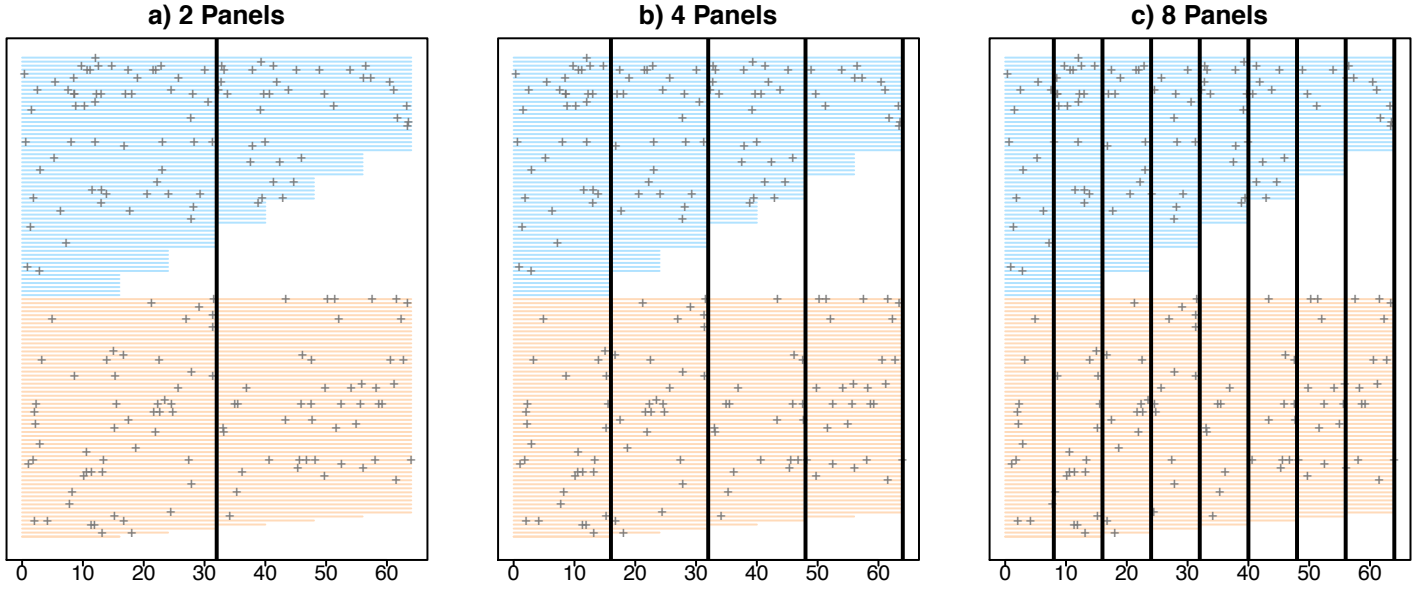


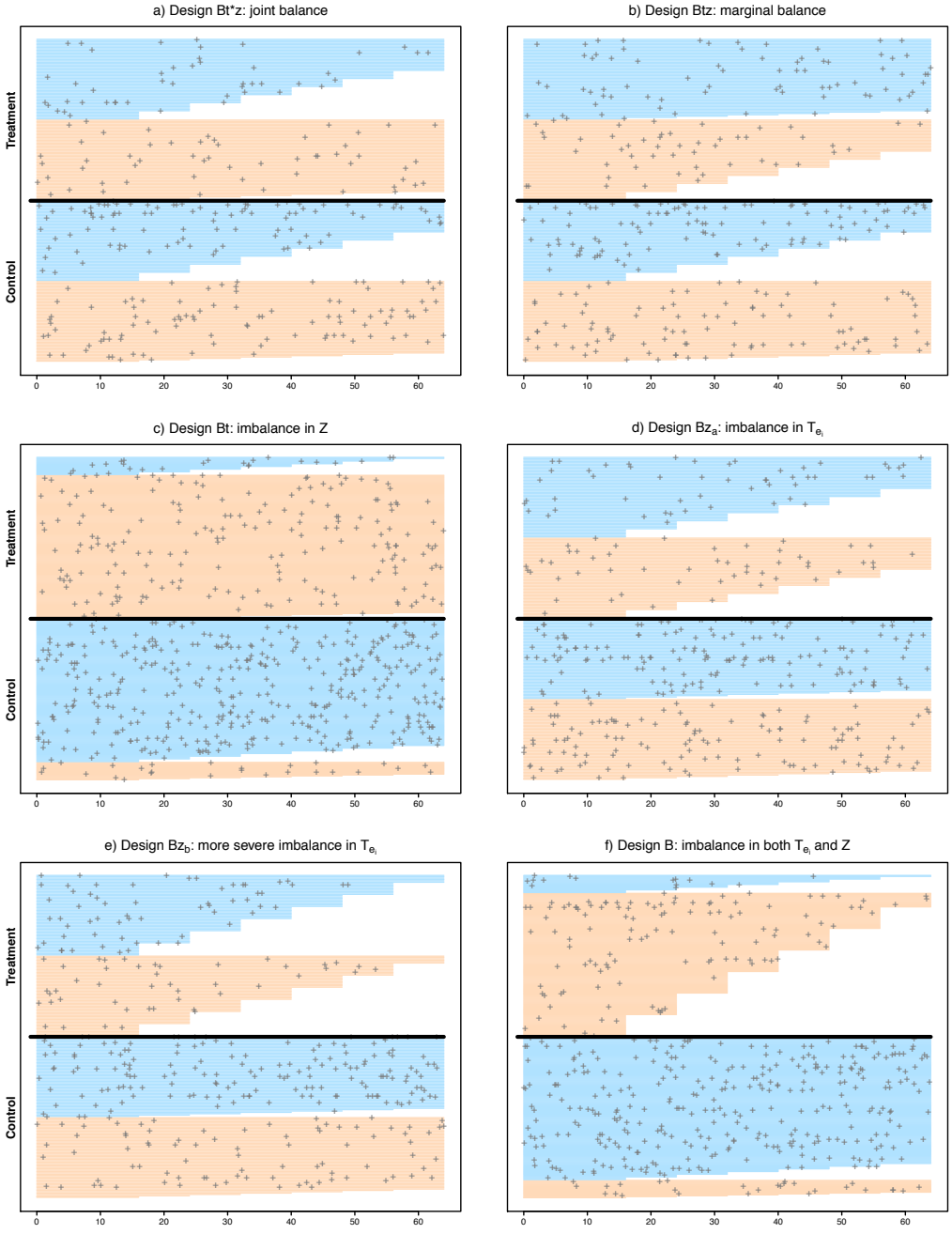
Table 2.1: Scenarios considered for illustration of AREs. The name of the scenario reflects the variable for which there is balance, and '*' denotes joint balance.

Design	Description
Bt*z	Reflects the situation of *joint balance* whereby for each stratum of the covariate, the set of termination times are identical in the treatment and control groups. This yields a fully efficient estimator of the treatment effect when $\tau = 0$.
Btz	Refers to a situation where both (i) the termination times in each of the treatment and control groups are identical, but not within covariate strata (see Bt below), and (ii) the number of individuals within each covariate stratum are equal, (see Bz below).
Bt	Refers to a situation where joint balance does not occur, but the termination times are identical in the two treatment groups. The proportion of individuals with each termination time, ignoring covariate stratum, in the control and treatment groups are the same. In the example, most males are assigned to the control and most females to the treatment group.
Bz_a	Reflects the situation where there are equal numbers of individuals in each covariate stratum over the treatment groups, but the termination times are not balanced over treatment groups. In the example, the termination times are longer in the control group.
Bz_b	Reflects a more extreme situation of imbalance than Bz_a.
B	Neither joint nor any marginal balance.

Table 2.2: Scenarios considered for illustration of AREs. The name of the scenario reflects the variable for which there is balance; '*' denotes joint balance. A description of the scenarios is provided in Table 2.1.

		Number of individuals (Female | Male)														Dropout (%)		Balance		Joint
Design	Group	16	24	32	40	48	56	64	16	24	32	40	48	56	64	Female	Male	T_{e_i}	Z	Balance
Bt*z	Control	1	1	1	1	1	1	54	6	6	6	6	6	6	24	10	60	√	√	√
	Treat	1	1	1	1	1	1	54	6	6	6	6	6	6	24	10	60			
Btz	Control	1	1	1	1	1	1	54	6	6	6	6	6	6	24	10	60	√	√	×
	Treat	6	6	6	6	6	6	24	1	1	1	1	1	1	54	60	10			
Bt	Control	1	1	1	1	1	1	14	3	3	3	3	3	3	142	30	11	√	×	×
	Treat	1	1	1	1	1	1	154	3	3	3	3	3	3	2	3	90			
Bz_a	Control	1	1	1	1	1	1	54	1	1	1	1	1	1	54	10	10	×	√	×
	Treat	6	6	6	6	6	6	24	6	6	6	6	6	6	24	60	60			
Bz_b	Control	1	1	1	1	1	1	54	1	1	1	1	1	1	54	10	10	×	√	×
	Treat	9	9	9	9	9	9	6	9	9	9	9	9	9	6	90	90			
B	Control	1	1	1	1	1	1	14	3	3	3	3	3	3	142	30	11	×	×	×
	Treat	24	24	24	24	24	24	16	3	3	3	3	3	3	2	90	90			

Figure 2.3: Design scenarios for exploration of ARES. Females are represented in the lighter colour (salmon), and males in the darker colour (blue). Table 2.1 contains a detailed description of these scenarios, and Table 2.2 provides the definition of the design.



Though we investigated several parameter values, we present here results corresponding to values which are close to estimates from the analysis of the bladder cancer data discussed in the next chapter: $\alpha = 1$, α being the shape parameter of the Weibull distribution; $\beta_1 = -3.5$, and $\beta_2 = -0.5$ for the baseline and treatment effects, respectively; τ ranging between 0 and 2, reflecting no, and quite severe, overdispersion; and γ varying between -4 and 4 reflecting a very wide range of values of the covariate effect.

Figure 2.4 illustrates the asymptotic relative efficiency of the estimate of the treatment effect, $\tilde{\beta}_2$, for all the imbalanced designs listed in Table 2.2 and for a range of values of τ and γ. Generally, all the designs yield fairly high efficiencies (greater than .96). Recall that the design Bt*z yields full efficiency for $\tilde{\beta}_2$ when $\tau = 0$; for all parameter settings when $\tau > 0$, efficiencies were greater than 0.99, and so are not displayed here. Design Btz, with marginal balance for both T_{e_i}'s and the covariates, also yields very high efficiencies of at least 0.98 for $\tilde{\beta}_2$. The same is true for design Bt, even with its high imbalance in the covariates. The design Bz_a, with marginal balance for the covariates and imbalance for the T_{e_i}'s also yields efficiencies of at least 0.98. Designs Bz_b and B yield slightly lower efficiencies; the most extreme case of imbalance over T_{e_i}'s, occurring in design Bz_b, yields efficiencies of 0.96 for the 2-panel study.

Figures 2.5 and 2.6 present the efficiencies of the estimators of the parameters for the overall mean process, i.e. the parameters α and β_1. Generally AREs for these parameters are much lower than those for the treatment estimator. The overall mean level of the intensity function, reflected in β_1, seems to be reasonably well estimated when there are at least 4 panels, whereas at least 8 panels are required to estimate the intensity shape parameter.

In summary, the efficiency of the estimate of the treatment effect is generally very high. Especially if there is balance in the T_{e_i}'s, even with extreme imbalance in the covariates, the efficiency will remain very high. Estimating the mean can be reasonably well achieved with 4 panel followup times; but estimating the shape parameter of the intensity function requires a greater number of panel followup times; we recommend at least 8.

Figure 2.4: AREs of the estimate of the treatment effect, $\tilde{\beta}_2$, for different designs with 2, 4, or 8 panels, varying values of the covariate effect, γ (x-axis), and the overdispersion parameter, τ (y-axis).

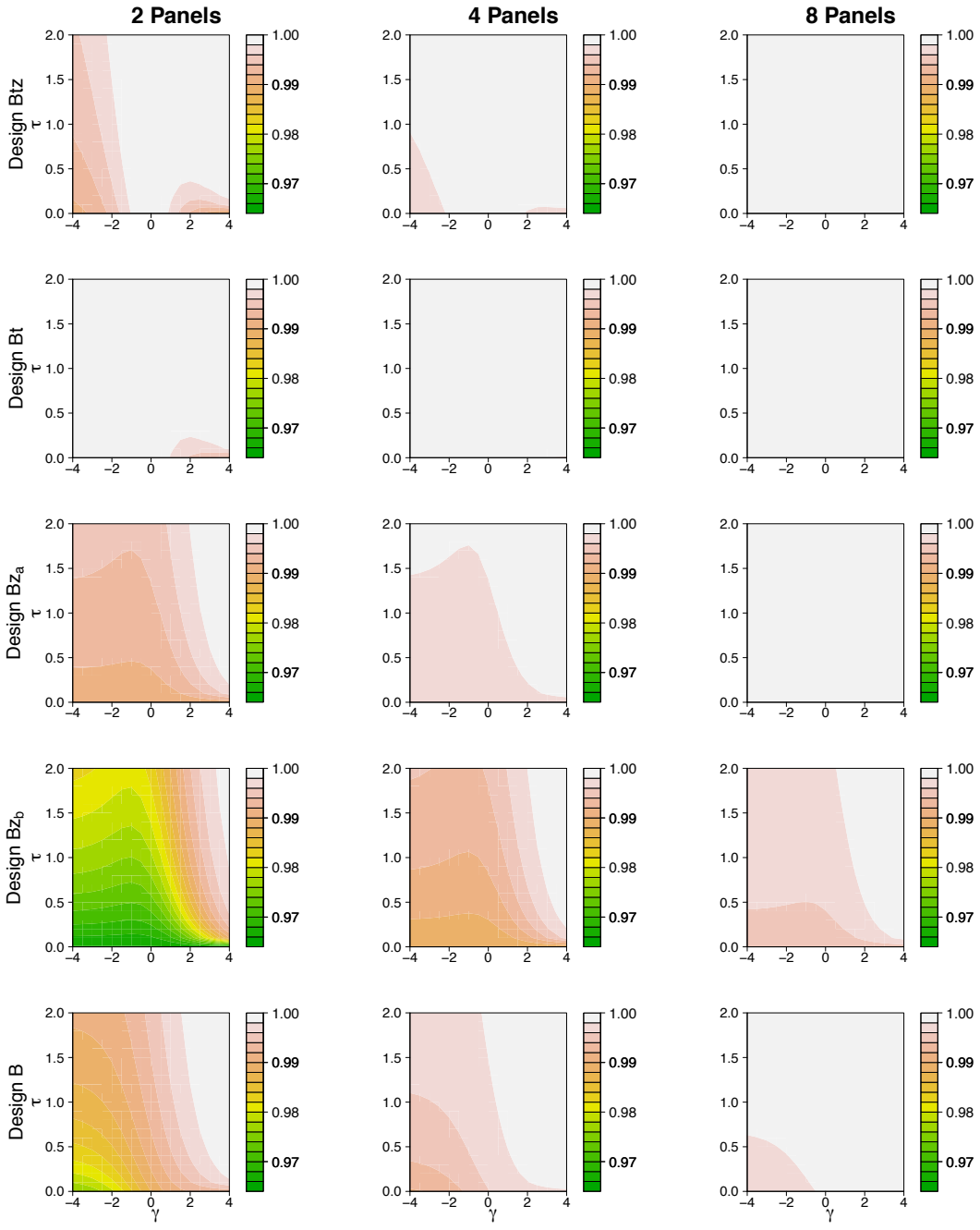


Figure 2.5: AREs of the estimator of α, the shape parameter of the baseline intensity function, for different designs with 2, 4, and 8 panels, varying values of the covariate effect, γ (x-axis), and of the overdispersion parameter, τ (y-axis).

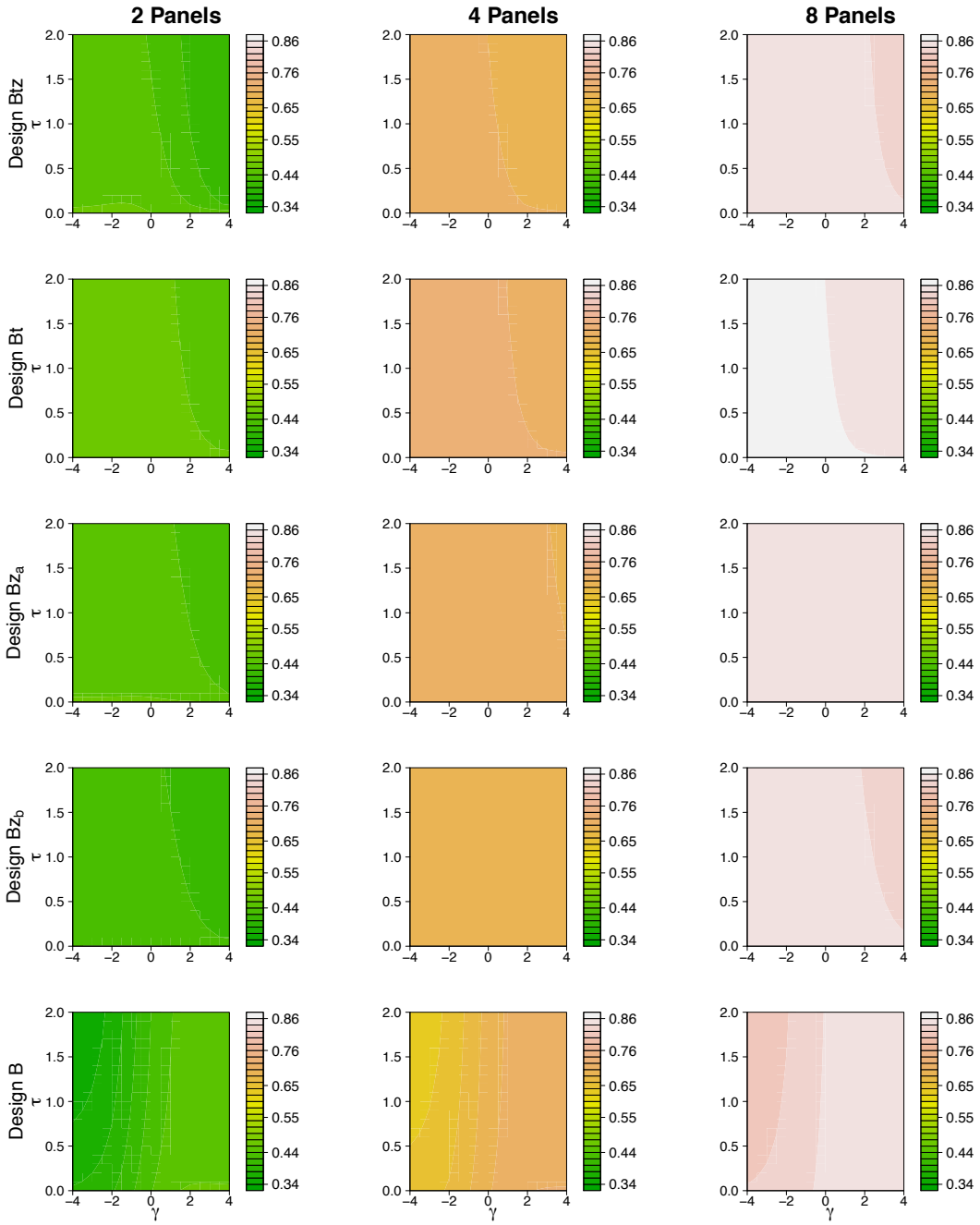
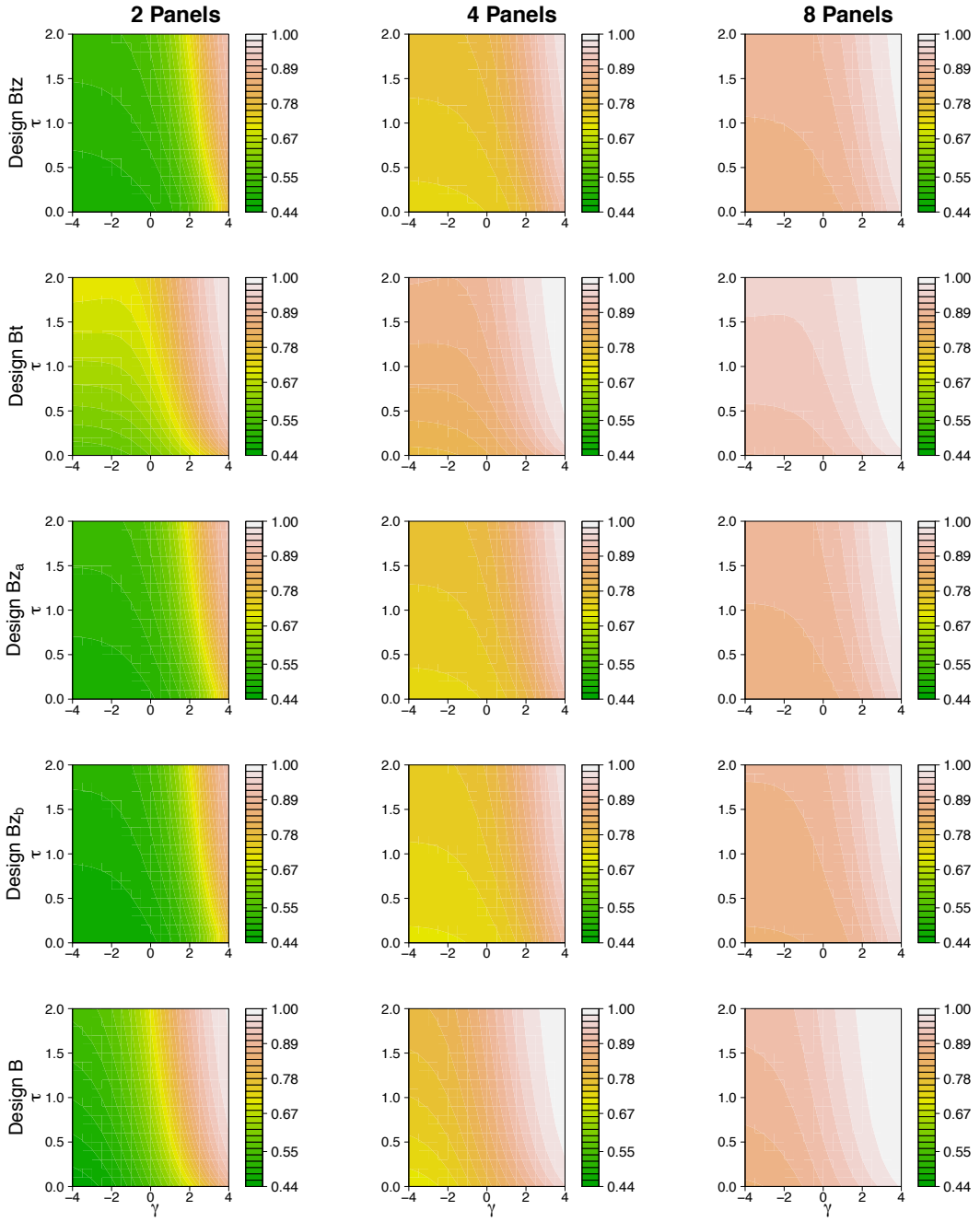


Figure 2.6: AREs of the estimator of β_1, the overall mean baseline treatment effect, for different designs with 2, 4, and 8 panels, varying values of the covariate effect (x-axis), γ, and of the overdispersion parameter, τ (y-axis).



2.3.2 Efficiency under imbalanced designs: 2-segments

We consider the designs discussed in Section 2.3.1 under three different situations where the segments are induced by a cut point at 16, 24 and 32 months, and where there are 1 and 2 panels per segment; we present the AREs of the estimates of the segment-specific and time-weighted mean treatment effects in Figures 2.7, 2.8, and 2.9. Recall that for the design with *joint balance*, Bt*z, the efficiency is 1 when $\tau = 0$; even in the presence of overdispersion, this design provides nearly fully efficient estimators of the treatment effect. The values of the segment-specific treatment effects, used to obtain the AREs when segment cutpoints are changed, correspond to within-segment mean values of these effects from assumed continuous time-varying functional forms which are decreasing/increasing over time for β_2^1/β_2^2.

Regarding imbalanced designs, efficiencies of the estimators of the treatment effects in segment 1 are high ($>$0.90) since the termination times in segment 1 are almost balanced (Figure 2.7). Efficiencies of the treatment estimators in segment 2, on the other hand, are lower than those in segment 1, with values as low as 0.65, but they increase as the cut point for the segment is shifted closer to the end of the study (Figure 2.8). With such a shift, the imbalance in the termination times is shared over the two segments, rather than being concentrated in segment 2; as well, the efficiency of the estimator of the treatment effect in segment 1 decreases only slightly.

The efficiency of the estimator of the time-weighted mean treatment effect is mostly driven by the level of balance in segment 2 because the estimator of the treatment effect in segment 1 tends to have high efficiency no matter the placement of the cut point defining the segments (Figure 2.9). Efficiencies of the estimators of segment-specific effects and time-weighted means increase as the number of panels increases.

Figure 2.7: AREs of $\tilde{\beta}_2^1$ for the 2-segment model for different designs with 1 and 2 panels per segment. Parameter values are $\alpha = 1.1$, $\beta_1 = -3.5$ for the baseline intensity function; $\beta_2^1 = 0.2, 0.075, -0.050$ when the segment cut point is at 16, 24, and 32 months, respectively; $\beta_2^2 = -0.8, -0.925, -1.050$ when the segment cut point is at 16, 24, and 32 months, respectively.

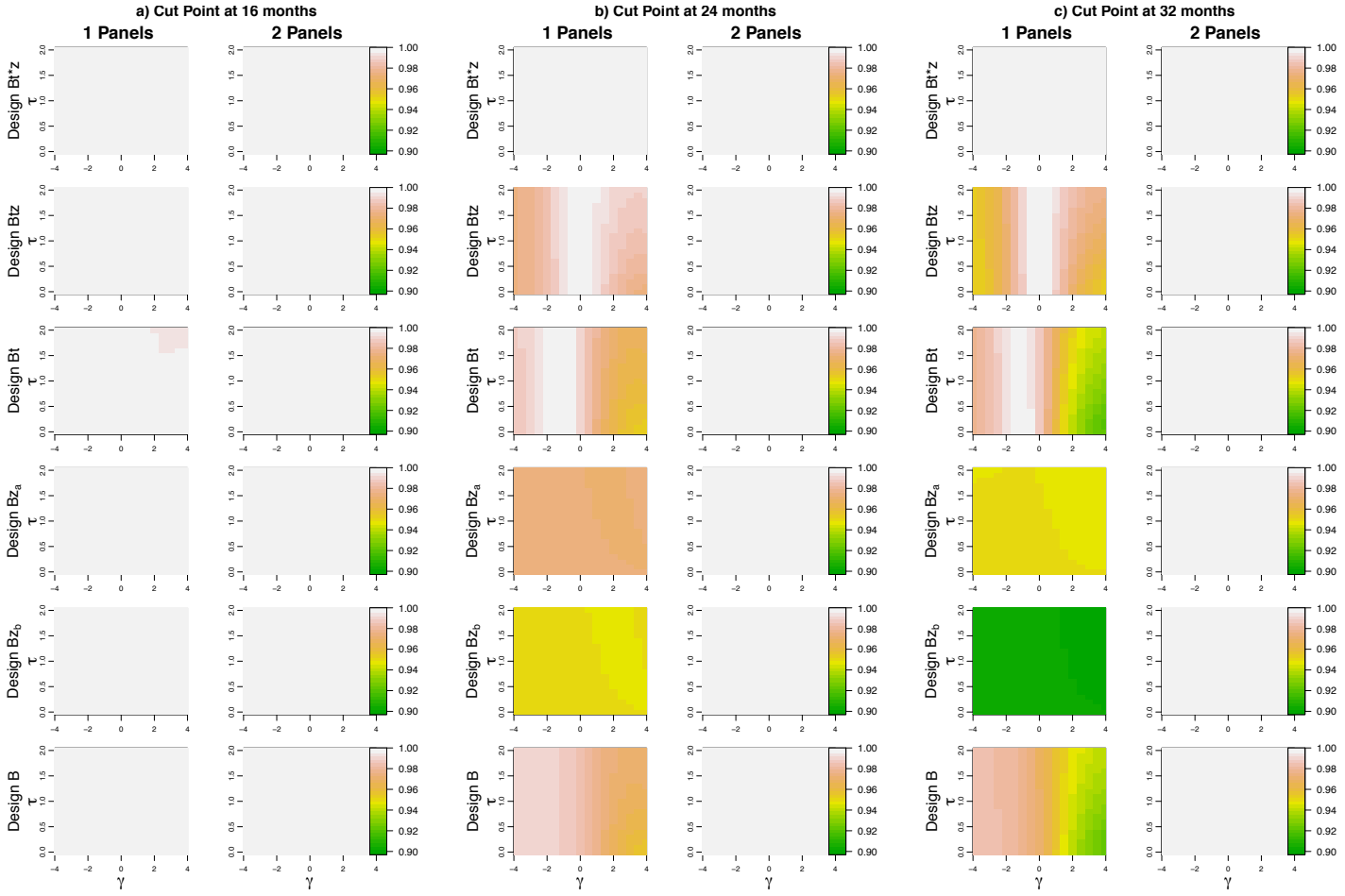


Figure 2.8: AREs of $\tilde{\beta}_2^2$ for the 2-segment model for different designs with 1 and 2 panels per segment. Parameter values are $\alpha = 1.1$, $\beta_1 = -3.5$ for the baseline intensity function; $\beta_2^1 = 0.2, 0.075, -0.050$ when the segment cut point is at 16, 24, and 32 months, respectively; $\beta_2^2 = -0.8, -0.925, -1.050$ when the segment cut point is at 16, 24, and 32 months, respectively.

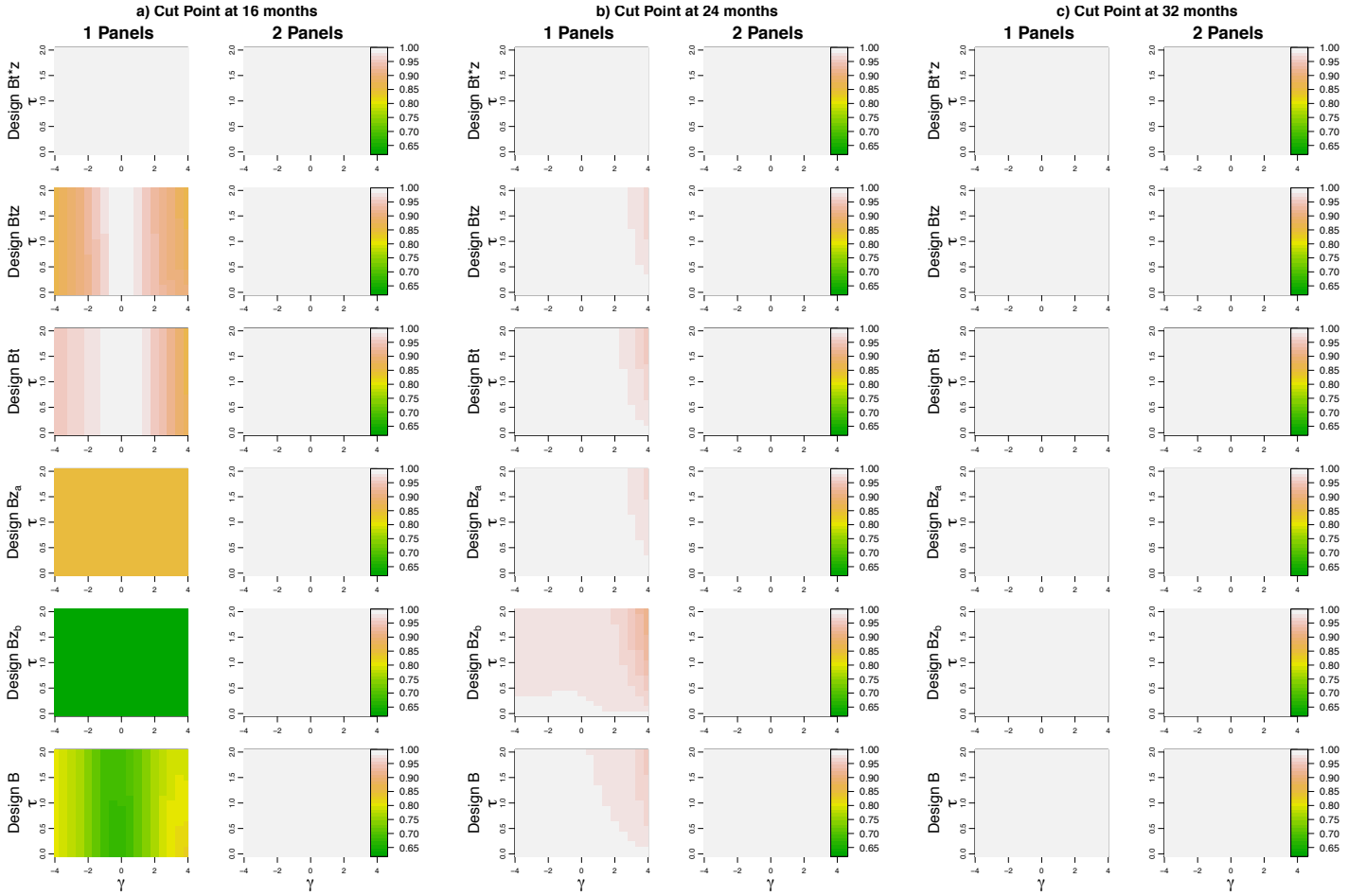
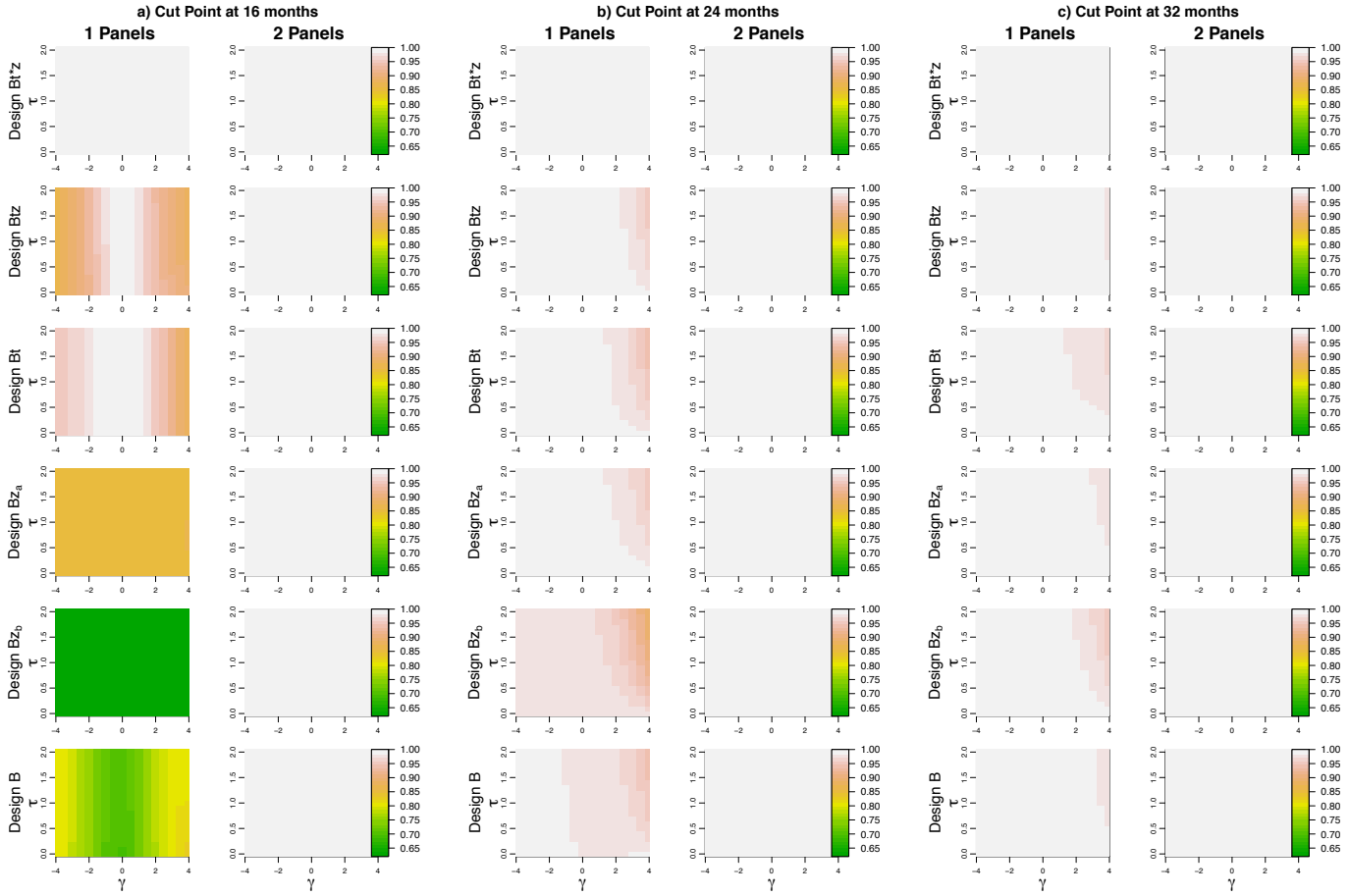


Figure 2.9: AREs of the estimate of the time-weighted mean treatment effect, $\tilde{\delta}_2$, based the 2-segment model for different designs with 1 and 2 panels per segment. Parameter values are $\alpha = 1.1$, $\beta_1 = -3.5$ for the baseline intensity function; $\beta_2^1 = 0.2, 0.075, -0.050$ when the segment cut point is at 16, 24, and 32 months, respectively; $\beta_2^2 = -0.8, -0.925, -1.050$ when the segment cut point is at 16, 24, and 32 months, respectively.



2.4 Discussion

In this investigation of the efficiency of panel designs relative to continuous followup for the study of recurrent events, we have focused on the estimation of a treatment effect in a semiparametric proportional intensity model. We have provided specific conditions for full efficiency, which relate principally to *joint balance* in the distribution of the termination times and covariates in the baseline and treatment groups. We note that imbalance in the distribution of the covariates over treatment is less detrimental in terms of sharply reducing efficiencies than imbalance in the termination times. With panel designs of at least 2 panels, the efficiency of the estimator of the treatment effect is high, and those with 4 to 8 panels give reasonable estimation of the shape of the baseline intensity function.

Detailed conditions for efficient study designs are provided as remarks in Section 2.2. The investigation of these conditions provides a basis for understanding the utility of panel studies; further, the measures of balance offer direction on how to design panel studies so that they are most effective.

These efficiencies may also be utilized in straightforward ways when considering sample size of panel studies, as will be illustrated later in Section 3.3.

Finally, we close with the observation that even though we have considered extreme imbalance for investigative purposes, any marked difference in the distribution of termination times must be carefully investigated for signs of dependency between the observation process and the recurrent event process.

2.5 Proof of Theorem 1

Under standard conditions for the application of asymptotic results to estimating equations, $\sqrt{M}(\hat{\theta}-\theta)$ is asymptotically normal with asymptotic covariance

$$E\left(-\lim_{M\to\infty}\frac{\partial g_f}{\partial\theta}\right)^{-1}E\{\lim_{M\to\infty}g_f g_f'\}\left\{E\left(-\lim_{M\to\infty}\frac{\partial g_f}{\partial\theta}\right)^{-1}\right\}'. \tag{2.22}$$

The asymptotic variance of $\sqrt{M}(\hat{\eta}-\eta)$, $\eta=(\beta',\gamma',\alpha')'$, from the full data analysis is $(\lim_{M\to\infty}\frac{1}{M}I_f)^{-1}$, where I_f has the form (2.10), because of three sets of identities:

(i)$E(-\partial g_{\eta_1}/\partial\tau)=0$, (ii)$E(-\partial g_{\alpha,f}/\partial\tau)=\mathbf{0}$, and (iii) the $\{k+d_z+d_\alpha\}\times\{k+d_z+d_\alpha\}$ upper left submatrix of $E(g_f g_f')$ is the same as the corresponding submatrix of $E(-\partial g_f/\partial\theta)$.

Finite sample variance estimates are obtained by substituting $\hat{\theta}$ for θ and omitting the expressions $\lim_{M\to\infty}$. In this case there are two usual options for approximating the expectation of the terms in (2.22). The first is a model-based approach, which requires specification of 3rd and 4th moments and is used here to derive the results of Theorem 1. The second option, the one we have employed in the illustration, is an empirical approach, which substitutes $E\{\sum_{i=1}^M g_{if}g_{if}'\}$ by $\{\sum_{i=1}^M g_{if}g_{if}'\}$; where g_{if} denotes the contribution to the score equation from individual i.

Note

$$E\{-\frac{\partial g_f}{\partial\theta}\}=\begin{pmatrix} I_f & \mathbf{0} \\ \mathbf{b}' & b_0 \end{pmatrix},$$

and

$$E\{g_f g_f'\}=\begin{pmatrix} I_f & \mathbf{c} \\ \mathbf{c}' & c_0 \end{pmatrix}.$$

Here I_f is given in (2.10), $\mathbf{0}$, $\mathbf{b}$, and $\mathbf{c}$ are of dimension $(k+d_z+d_\alpha)\times 1$ vectors, $\mathbf{0}$ is a vector of zeros, $\mathbf{b}$ and $\mathbf{c}$ have elements

$$\begin{aligned} b_r &= \sum_{i=1}^M \frac{(1+2\tau\mu_i)y_{ir}}{(1+\tau\mu_i)}, \quad r=1,\ldots,k+d_z+d_\alpha, \\ c_r &= \sum_{i=1}^M \frac{\gamma_{3i}y_{ir}}{(1+\tau\mu_i)^3}, \quad \gamma_{3i}=E(Y_i-\mu_i)^3, \quad r=1,\ldots,k+d_z+d_\alpha, \end{aligned}$$

and

$$\begin{aligned} b_0 &= \sum_{i=1}^M \frac{\mu_i^2}{(1+\tau\mu_i)^2}, \\ c_0 &= \sum_{i=1}^M \frac{\gamma_{4i}-\mu_i^2(1+\tau\mu_i)^2}{(1+\tau\mu_i)^4}, \quad \gamma_{4i}=E(Y_i-\mu_i)^4. \end{aligned}$$

The asymptotic variance of $\hat{\theta}$ is then estimated as

$$\begin{pmatrix} I_f^{-1} & \frac{1}{b_0}I_f^{-1}\mathbf{c}-\mathbf{b} \\ \frac{1}{b_0}\mathbf{c}-\mathbf{b}'I_f^{-1} & \frac{1}{b_0^2}(c_0-2\mathbf{c}'I_f^{-1}\mathbf{b}+\mathbf{b}'I_f^{-1}\mathbf{b}) \end{pmatrix},$$

replacing θ by $\hat{\theta}$.

The asymptotic model-based variance of $\tilde{\theta}$ is estimated as above, replacing I_f with I_p and θ with $\tilde{\theta}$.

If we assume 3th and 4th moments as for the negative binomial distribution, $\gamma_{3i} = \mu_i(1+\tau\mu_i)(1+2\tau\mu_i)$, $\gamma_{4i} = 6\tau\mu_i^2(1+\tau\mu_i)^2 + \mu_i(1+\tau\mu_i)(1+3\mu_i+3\tau\mu_i^2)$, and $\mathbf{c} = \mathbf{b}$. In this case, the estimators of $\eta = (\beta', \gamma', \alpha')'$ and τ from either the full or panel data analysis are asymptotically independent. Note that only mean and variance assumptions are required for consistency of the asymptotic variance of $\hat{\eta}$ or $\tilde{\eta}$.

To obtain the asymptotic variances of $\tilde{\beta}$ and $\hat{\beta}$, Asvar$(\tilde{\beta})$ and Asvar$(\hat{\beta})$, respectively, we consider the partition of the information matrix (2.10) into blocks related to X, Z, and W, and obtain the inverse; these expressions are given by (2.23) and (2.24) below:

$$\text{Asvar}(\tilde{\beta}) = \tag{2.23}$$
$$(X'UX)^{-1} + (X'UX)^{-1}X'UZ\phi_{z,z}^{-1}Z'UX(X'UX)^{-1} + L_J(E+H_p)^{-1}L_J;$$

$$\text{Asvar}(\hat{\beta}) = \tag{2.24}$$
$$(X'UX)^{-1} + (X'UX)^{-1}X'UZ\phi_{z,z}^{-1}Z'UX(X'UX)^{-1} + L_J(E+H_f)^{-1}L_J;$$

where

$$\phi_{z,z} = Z'UZ - Z'UX(X'UX)^{-1}X'UZ,$$
$$E = \phi_{w,w} - \phi_{z,w}'\phi_{z,z}^{-1}\phi_{z,w},$$
$$\phi_{w,w} = W'UW - W'UX(X'UX)^{-1}X'UW,$$

$\phi_{z,w} = Z'UW - Z'UX(X'UX)^{-1}X'UW$, and

$$L_J = (X'UX)^{-1}X'UW - (X'UX)^{-1}X'UZ\phi_{z,z}^{-1}\phi_{z,w}. \tag{2.25}$$

The group-specific weighted covariation $\phi_{z,w}$ is given explicitly in (2.26):

$$\phi_{z,w(p,s)} = \sum_{j=1}^{k}\sum_{i\in G_j} u_i\left(z_{ip} - \frac{\left[uz_p\right]_{j+}}{\left[u\right]_{j+}}\right)\left(w_{is} - \frac{\left[uw_s\right]_{j+}}{\left[u\right]_{j+}}\right), \tag{2.26}$$

where $p = 1,2,\ldots,d_z$ and $s = 1,2,\ldots,d_\alpha$.

The elements of $(X'UX)^{-1}X'UZ$ are:

$$\begin{pmatrix} [u\mathbf{z}]_{1+}/[u]_{1+} \\ [u\mathbf{z}]_{2+}/[u]_{2+} - [u\mathbf{z}]_{1+}/[u]_{1+} \\ \vdots \\ [u\mathbf{z}]_{k+}/[u]_{k+} - [u\mathbf{z}]_{1+}/[u]_{1+} \end{pmatrix}, \tag{2.27}$$

and do not depend on the β_j's; they are function of the T_{e_i}'s and α and γ. Similar expressions may be obtained for $(X'UX)^{-1}X'UW$.

When α is a scalar, L_J simplifies to a $k \times 1$ vector (2.28):

$$L_J = (l_1, l_2, \ldots, l_k)' = (l_{w,1}, l_{w,2}, \ldots, l_{w,k})' - (l_{z,1}, l_{z,2}, \ldots, l_{z,k})' \frac{\phi_{z,w}}{\phi_{z,z}} \tag{2.28}$$

where

$$(l_{w,1}, l_{w,2}, \ldots, l_{w,k})' = (X'UX)^{-1}X'UW, \quad \text{and}$$
$$(l_{z,1}, l_{z,2}, \ldots, l_{z,k})' = (X'UX)^{-1}X'UZ.$$

Hence the asymptotic variances of the estimators $\tilde{\beta}$ and $\tilde{\alpha}$ are

$$\begin{aligned} \text{Asvar}(\tilde{\beta}_1) &= [u]_{1+}^{-1} + l_{z,1}^2\phi_{z,z}^{-1} + l_1^2(H_p + E)^{-1}, \\ \text{Asvar}(\tilde{\beta}_j) &= [u]_{j+}^{-1} + [u]_{1+}^{-1} + l_{z,j}^2\phi_{z,z}^{-1} + l_j^2(H_p + E)^{-1}, j = 2,3,\ldots,k, \\ \text{and} & \\ \text{Asvar}(\tilde{\alpha}) &= (H_p + E)^{-1}. \end{aligned}$$

The inverse of I_f, the information matrix based on the full data, can be similarly computed to obtain $\text{Asvar}(\hat{\beta}_j)$, $j = 1,\ldots,k$, and $\text{Asvar}(\hat{\alpha})$. These are calculated using identical formulae as above, except H_p is replaced with H_f. Thus, Theorem 1 follows from computing the ratio of these asymptotic variances, and the precise formulation stated offers emphasis on the important elements regarding efficiency.

2.6 Notation

i: indexes individuals

t: indexes time

k: number of treatments under test

s: indexes segments; $s = 1, \ldots, S$

m_j: number of individuals on treatment j, $j = 1, \ldots, k$

M: sample size of the study; $M = \sum_{j=1}^{k} m_j$

$Y_i(t)$: observation process for individual i

$N_i(t)$: counting process for individual i

$\bar{N}_i(t)$: observed counting process for individual i; $\bar{N}_i(t) = \int_0^t Y_i(u) dN_i(u)$

x_i: vector of treatment indicators for individual i

z_i: covariates for individual i

ν_i: individual-specific random effect $i = 1, \ldots, M$

α: parameters determining the baseline intensity function

β: treatment effects

γ: covariate effects

η_1: regression parameters; $\eta_1 = (\beta', \gamma')'$

η: regression parameters including parameters in the baseline intensity; $\eta = (\beta', \gamma', \alpha')'$

τ: overdispersion parameter

θ: full set of parameters; $\theta = (\beta', \gamma', \alpha', \tau)'$

$(T^{s-1}, T^s]$: time period of segment s; $T^0 = 0$

$I_{(T^{s-1},T^s]}(t)$: indicator of whether individual i is observed at time t

$\lambda_i(t)$: counting process intensity function

- 1-segment: $\lambda_i(t) = \nu_i \rho(t;\alpha) e^{\mathbf{x}_i'\beta + \mathbf{z}_i'\gamma}$
- S segments: $\lambda_i(t) = \nu_i \rho(t;\alpha) \exp\left\{\mathbf{x}_i' \left[\sum_{s=1}^{S} \beta^s I_{(T^{s-1},T^s]}(t)\right] + \mathbf{z}_i'\gamma\right\}$

$\rho(t;\alpha)$: baseline intensity function

d_α: dimension of α

d_z: dimension of γ

a: indexes parameters in the baseline intensity; $a = 1,\ldots,d_\alpha$

n_{ip}: number of events observed in panel p for individual i, $p = 1,\ldots,e_i$

n_{i+}: total number of events observed for individual i; $n_{i+} = \sum_{p=1}^{e_i} n_{ip}$

$(T_{i,p-1}, T_{i,p}]$: pth panel period observed for individual i (for the 1-segment study)

T_{e_i}: termination for individual i (for the 1-segment study)

R_{ip}: cumulative baseline intensity function over panel period p for individual i (for the 1-segment study); $R_i = \int_{T_{i,p-1}}^{T_{i,p}} Y_i(t)\rho(t;\alpha)dt$

R_i: cumulative baseline intensity function over the entire observation period for individual i; $R_i = \int_0^\infty Y_i(t)\rho(t;\alpha)dt$

μ_{ip}: expected mean number of events for individual i in panel p (for the 1-segment study); $\mu_{ip} = R_{ip}\exp\{\mathbf{x}_i'\beta + \mathbf{z}_i'\gamma\}$

μ_{i+}: total expected mean number of events for individual i; $\mu_{i+} = R_i \exp\{\mathbf{x}_i'\beta + \mathbf{z}_i'\gamma\}$

L_p: likelihood based on panel data

L_f: likelihood based on full data, i.e. continuous followup; $L_f = L_{\alpha,f} L$

$w_{ia} = \partial \log L_{\alpha,f}(\alpha)/\partial\alpha_a$: a function of the termination time for individual i

ω_{ipl}: time of the lth event from start time ($t = 0$) for individual i in panel period p

g_p: estimating equations for panel data analysis

g_f: estimating equations for full data analysis

g_{η_1}: estimating equations for $(\beta', \gamma')'$

$g_{\alpha,d}$: estimating equations for α

g_τ: estimating equation for τ

h_i: correction term to reduce small sample bias;
$h_i = \text{diag}(U_1^{1/2} V_1' (V_1' U_1 V_1)^{-1} V_1' U_1^{1/2})$

V: treatment and other covariates $V = (\ X \quad Z\ )$

W: matrix with entries w_{ia}

V_1: covariates including W; $V_1 = (\ X \quad Z \quad W\ )$

$u_i = \mu_{i+}/(1 + \tau\mu_{i+})$: function of the expected number of events and the overdispersion parameter for individual i

$T_{i,p}$: pth panel followup time for individual i, $i = 1, \ldots, M$, $p = 1, \ldots, e_i$

T_{e_i}: termination time for individual i; also denoted as T_{i,e_i}

I_p, I_f: information matrix based on likelihoods for the panel and full data, respectively

$\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\tau}$: estimators obtained from an analysis of panel data

$\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\tau}$: estimators obtained from an analysis of the full data

G_j: set of individuals who receive treatment j

$[n]_{j+}$: total number of events observed for all individuals receiving treatment j, $[n]_{j+} = \sum_{i \in G_j} n_{i+}$

$[u]_{j+}$: function of the expected number of events for all individuals receiving treatment j, $[u]_{j+} = \sum_{i \in G_j} u_i$

$[uz]_{j+}$: function of the expected number of events and the covariates for all individuals receiving treatment j, $[uz]_{j+} = \sum_{i \in G_j} u_i z_i$

$[uw]_{j+}$: function of the expected number of events and the termination times for all individuals receiving treatment j, $[uw]_{j+} = \sum_{i \in G_j} u_i w_i$

E: corrected variation of the w_i's, $E = \phi_{w,w}(1 - \phi_{z,w}^2/\phi_{z,z}\phi_{w,w})$

$\phi_{z,w}$: weighted covariation of z_i's and w_i's,
$\phi_{z,w} = \sum_{j=1}^{k} \sum_{i \in G_j} u_i \left(z_i - [uz]_{j+}/[u]_{j+}\right)\left(w_i - [uw]_{j+}/[u]_{j+}\right)$

$\phi_{z,z}$: weighted variation of z_i's, $\phi_{z,z} = \sum_{j=1}^{k} \sum_{i \in G_j} u_i \left(z_i - [uz]_{j+}/[u]_{j+}\right)^2$

$\phi_{w,w}$: weighted variation of w_i's, $\phi_{w,w} = \sum_{j=1}^{k} \sum_{i \in G_j} u_i \left(w_i - [uw]_{j+}/[u]_{j+}\right)^2$

$r_{z,w} = \phi_{z,w}/(\phi_{z,z}\phi_{w,w})$: ratio of weighted covariation between z_i and w_i; weights are u_i.

l_1: measure of balance relating to β_1, $l_1 = [uw]_{1+}/[u]_{1+} - [uz]_{1+}/[u]_{1+} \times \phi_{z,w}/\phi_{z,z}$,

l_j: measure of balance relating to β_j, $j \geq 2$;
$l_j = \sqrt{\phi_{w,w}}\left\{l_{w,j}/\sqrt{\phi_{w,w}} - l_{z,j}/\sqrt{\phi_{z,z}} r_{z,w}\right\}$

$l_{z,j}$: measure of balance relating to β_j, and with respect to covariates; $l_{z,j} = [uz]_{j+}/[u]_{j+} - [uz]_{1+}/[u]_{1+}$

$l_{w,j}$: measure of balance relating to β_j, and with respect to termination times; $l_{w,j} = [uw]_{j+}/[u]_{j+} - [uw]_{1+}/[u]_{1+}$

n_{ip}^s: number of observed events for individual i in panel p in segment s, $n_{ip}^s = \bar{N}_i(T_{i,p}^s) - \bar{N}_i(T_{i,p-1}^s)$

μ_{ip}^s: expected number of events for individual i in panel p in segment s, $\mu_{ip}^s = E(n_{ip}^s) = R_{ip}^s e^{\mathbf{x}_i' \beta^s + \mathbf{z}_i' \gamma}$

R_{ip}^s: cumulative baseline function for individual i in panel p in segment s,
$R_{ip}^s = R_{ip}^s(T_{i,p-1}^{s-1}, T_{i,p}^s; \alpha) = \int_{T_{i,p-1}^{s-1}}^{T_{i,p}^s} \rho(u;\alpha) Y_i(u) du$

μ_{i+}^s: expected number of events for individual i in segment s,
$\mu_{i+}^s = \sum_{p=1}^{e_i^s} \mu_{ip}^s = \sum_{p=1}^{e_i^s} R_i^s e^{\mathbf{x}_i' \beta^s + \mathbf{z}_i' \gamma}$

R_i^s: cumulative baseline function for individual i in segment s, $R_i^s = \sum_{p=1}^{e_i^s} R_{ip}^s$

u_i^s: $u_i^s = \mu_{i+}^s / (1 + \tau \mu_{i+}^s)$

w_i^s: function of the T_{e_i}'s for individual i in segment s, $w_i^s = \partial \log R_i^s / \partial \alpha$

β^s: treatment effects in segment s; β_j^s: effect of the jth treatment in segment s

$l_{z,j}^s, l_{w,j}^s, l_j^s$: segment-specific measures of balance with respect to covariates, termination times, and overall, respectively.

$[u]_{j+}^s$: function of the expected number of events for all individuals receiving treatment j in segment s, $[u]_{j+}^s = \sum_{i \in G_j} u_i^s$

$[uz]_{j+}^s$: function of the expected number of events and the covariates for all individuals receiving treatment j in segment s, $[uz]_{j+}^s = \sum_{i \in G_j} u_i^s z_i$

$[uw]_{j+}^s$: function of the expected number of events and the termination times for all individuals receiving treatment j in segment s, $[uw]_{j+}^s = \sum_{i \in G_j} u_i^s w_i^s$

Chapter 3

Illustration of Efficient Designs and Power Considerations

3.1 Introduction

In this chapter we illustrate the methods from Chapter 2 in an analysis of the recurrence of tumors in patients with bladder cancer (Andrews and Herzberg, 2000), and we present alternative uses of the methods developed which are helpful for power analysis and sample size calculations.

3.2 Bladder Cancer Data

A clinical trial, conducted by the Veterans Administrative Co-operative Urological Research Group (Byar et al., 1977), studied the effects of placebo pills, pyridoxine pills, and periodic instillation of thiotepa into the bladder on the frequency of recurrence of bladder cancer. The data appear in Andrews and Herzberg (2000). All 116 patients had bladder cancer when they entered the study; the tumors were removed and the patients were randomly assigned to one of the three treatments. We consider two covariates which may reflect cancer severity at baseline: the number of tumors and the size, in centimetres, of the largest tumor. Here we consider estimation of the treatment effect under a design with continuous followup and in artificial designs, for illustrative purposes, with 2, 4, and 8

equally spaced scheduled followup visits over 64 months; for the panel designs, we record information on event counts at the scheduled followup times and at termination times.

3.3 Measures of Balance

Because the balance in the covariates and the termination times T_{e_i}s play an important role in the efficiency of the estimator of the treatment effect, we provide a brief analysis of such balance before presenting the results of the analyses. We consider a modified version of the bladder cancer dataset in which we have truncated some termination times to achieve imbalance so as to assess its impact on estimation of the treatment effects. To create the modified data, 80% of the individuals in the placebo group with termination times over 24 months were randomly assigned new termination times generated from a Uniform(6,24) distribution. Figure 3.1 displays both datasets: original (left panel) and modified (right panel). Note that the rate of occurrence of events seems to be slightly lower in the thiotepa group. The more detailed comparisons of these data displayed in Figure 3.2 shows that in the original data the three treatment groups are very similar in terms of the distributions of the termination times, the number of tumors, and the size of the largest tumor at baseline. We thus expect the efficiency of the treatment effect estimator to be high in the analysis of the original data. On the other hand, in the modified dataset, because of shorter termination times in the placebo group, we expect somewhat lower efficiency of the estimator of the treatment effect (Figure 3.1 and Figure 3.2).

In addition to visual presentations of balance, we provide in Table 3.1 two types of measures of balance, (1) derived from estimates of parameters obtained by fitting the full data assuming continuous followup, model based measures, $l_{w,j}$, $l_{z,j}$, and $r_{z,w}$ (see Theorem 1 from Chapter 2), and (2) summary measures, l_j^* and l_j^{**}. The measure l_j^{**} is a raw measure of balance in the termination times. It compares the proportions of individuals with termination times 0-8, 8-16, 16-24, 24-32, 32-40, 40-48, 48-56, and 56-64 months in the placebo group with those in the treatment group j. Specifically, it is calculated as the sum of squared differences between these proportions for treatment group j and placebo group. Similarly, l_j^* is a raw summary of joint balance between the termination times and the number of tumors. It is computed as a sum of squared differences between the proportions of individuals in the placebo group and treatment group j falling into 16 categories

defined by their termination times (8 groups; as for l_j^{**}) and number of tumors (2 groups: 0 or 1 tumor, and 2 or more tumors). For our illustrative purposes we do not show here joint balance measures related to the size of the largest tumor as this covariate is not significant. The greater the imbalance, the larger all measures will be. Note that, for the original data, although $l_{z,j}$ is large, l_j, l_j^* and l_j^{**} are all small, indicating high balance in the data. Measures are larger for the modified data.

Table 3.1 presents the ratios H_p/H_f, which play a key role in the AREs of the estimates of the baseline intensity parameters. This ratio increases with the number of panels more steeply using the original, than using the modified, design. Table 3.1 also provides the corresponding measures of balance for the model where the treatment effect is piecewise constant over two equally spaced segments. Measures reveal that the overall mean treatment effects are expected to be less efficient in the analysis of the modified data since the ratio of the segment specific expected number of events is almost the same for the two datasets (original and modified), but l_j is considerably lower in the original dataset. Similar raw measures of balance as l_j^*, and l_j^{**} computed for the 2-segment model (not shown here) also indicate large differences in balance between the original and modified data. We also note that the modified data appears quite balanced over segment one and imbalanced over segment 2, and this difference will be reflected in standard error estimates corresponding to these segments (illustrated in Section 3.4).

Figure 3.1: Bladder cancer data: the plot displays the termination times for the original data (a), and modified data (b). For the display of the covariates and individual-specific rates, the length of the lines are proportional to the size of the largest tumor, the number of tumors, and the rate of events, respectively; additionally, lighter colors are used to highlight large values of these variables.

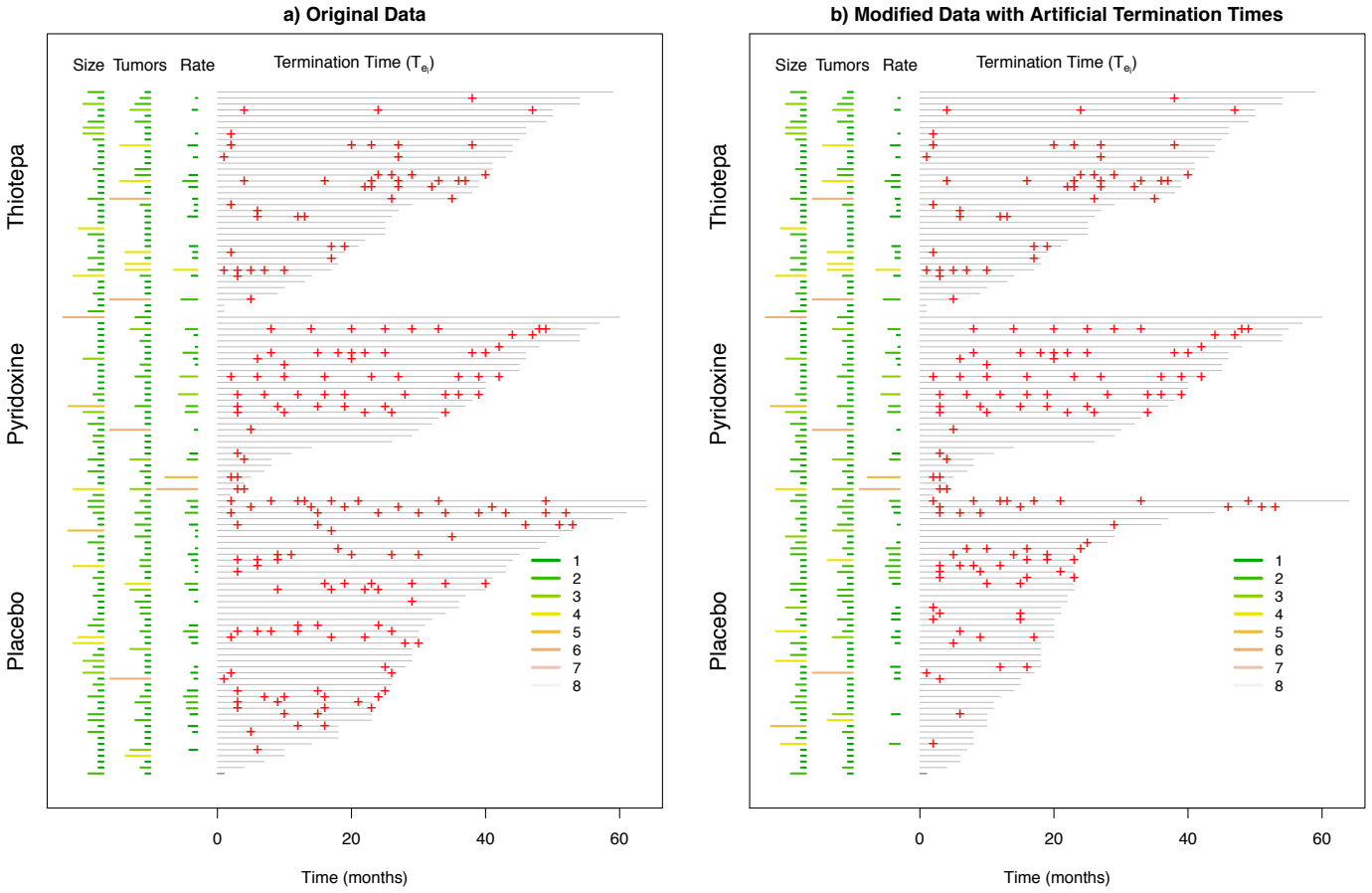


Figure 3.2: Distributions of variables in the bladder cancer data: original data (a), and modified data (b). The smooth line corresponds to a smoother based on a Gaussian kernel. The mean and the median are displayed on the plots.

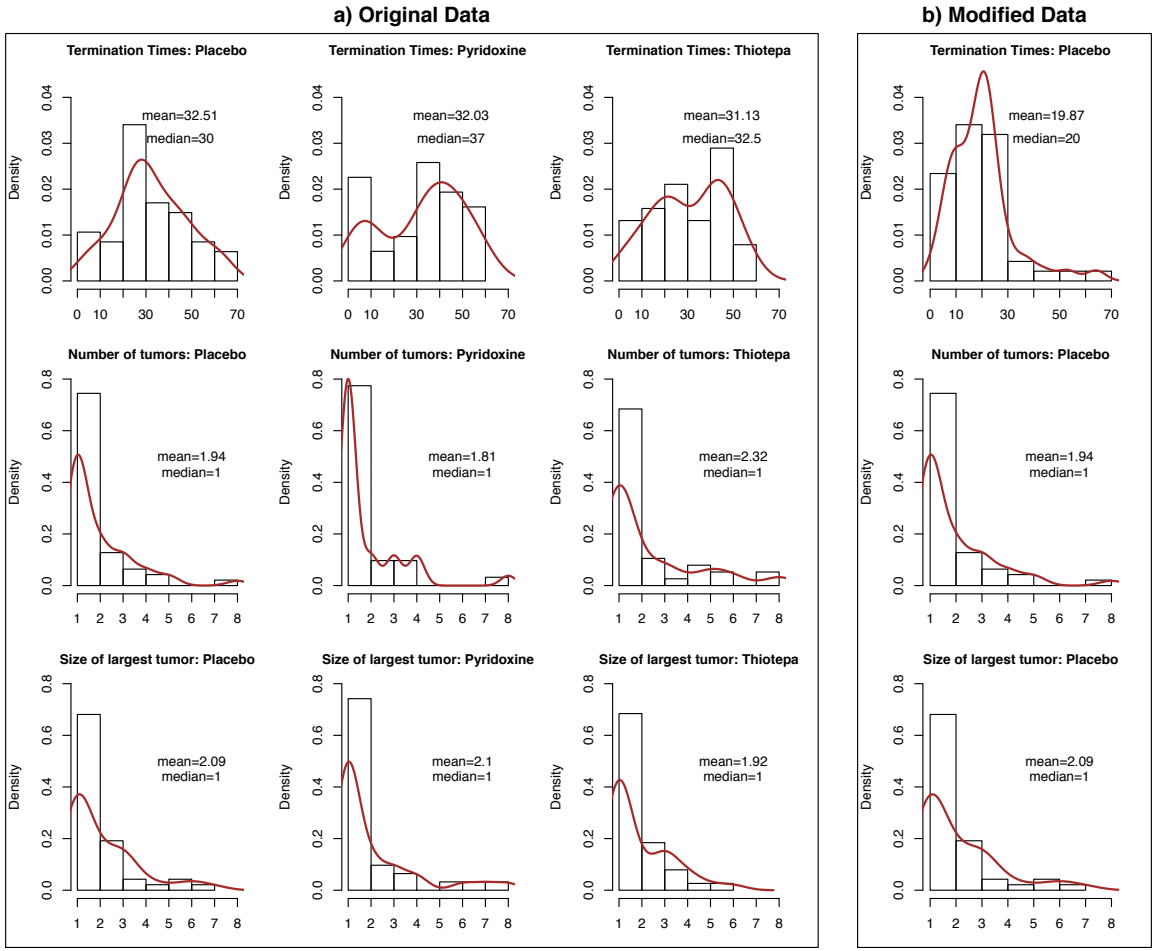


Table 3.1: Measures of balance in the bladder cancer data. The weighted differences of the w_i's, $l_{w,j}$, the weighted differences of z_i's, $l_{z,j}$, and the overall measure of balance l_j correspond to model-based measures; l_j^{**} and l_j^* correspond to the sum of squared differences of proportions of individuals per strata in the two treatment groups. For the 2-segments case, l_j^1 and l_j^2 correspond to measures of balance in segment 1 and 2, respectively; l_j corresponds to the measure of balance for the time-weighted overall mean treatment effect. Note that $r_{z,w}$ and the ratios H_p/H_f do not vary by treatment group, so they are displayed only once; 2p, 4p, and 8p denote 2, 4, and 8 panels, respectively.

	Model-based measures				Raw measures		H_p/H_f		
	1-segment Measures								
	$l_{w,j}$	$l_{z,j}$	l_j	$r_{z,w}$	l_j^{**}	l_j^*	2p	4p	8p
	Original Data								
Pyridoxine	0.004	-0.130	-0.001	-0.110	0.083	0.068	0.175	0.513	0.717
Thiotepa	-0.010	0.668	0.013		0.028	0.029			
	Modified Data								
Pyridoxine	0.464	-0.287	0.459	-0.059	0.297	0.167	0.164	0.454	0.668
Thiotepa	0.453	0.516	0.463		0.175	0.104			
	2-segment Measures								
	l_j^1	l_j^2	l_j				2p	4p	8p
	Original Data								
Pyridoxine	0.394	-1.170	-0.388				0.000	0.393	0.645
Thiotepa	0.367	-1.176	-0.405						
	Modified Data								
Pyridoxine	-0.128	-1.722	-0.925				0.000	0.339	0.595
Thiotepa	-0.147	-1.712	-0.929						

3.4 Analysis of Bladder Cancer Data

Table 3.2 reports parameter estimates and their standard errors based on the modified data under a 2-panel, a 4-panel, and an 8-panel design, as well as an analysis of the full data; Table 3.3 reports results from corresponding analyses for the original data. Both tables present results from two analyses: one that considers 1 segment, so the treatment effects are fixed, and another that considers 2 segments, with treatment effects being piecewise constant over segments of 32 months. Since the estimated coefficient corresponding to the size of the largest tumor at baseline is not significant, this variable has been excluded in the subsequent discussion. Because there is close to *joint balance* in the original data, the treatment estimators from the three panel designs are almost fully efficient compared to the analysis of the full data (Table 3.3). On the other hand, in the imbalanced data, the treatment estimators from the 2-panel design have slightly higher standard errors than those obtained from the full data; the efficiency is quickly recovered by the 4-panel design. The thiotepa treatment, β_3=-0.485 (Std. Error =0.281), may have a protective effect on recurrences relative to placebo, while the pyridoxine treatment effect is non-significant. Note that for the model where we consider 2 segments, the effect of thiotepa is more protective in the second segment. The estimates of the time-weighted mean treatment effects in the 2-segment model, $\tilde{\delta_2}$ and $\tilde{\delta_3}$, have standard errors which are slightly higher in the 2-panel design than in the 8-panel and full data designs. There is substantial overdispersion in the data, and the estimate of the Weibull shape parameter α is quite close to unity. The estimate of the standard error of $\tilde{\alpha}$ is considerably higher in the 2-panel design, but much closer in the 8-panel design, than that of the corresponding estimate from the full data analysis.

3.5 Simulation: Power Considerations

Consider the construction of a 1-segment panel design generating current status data in a recurrent event study intended to estimate a treatment effect. Current status data are commonly encountered in epidemiological studies where information on the number of events is only available at the end of the study (Sun and Kalbfleisch, 1993); the data structure also arises in carcinogenicity experiments where the number of tumors is only known at sacrifice of the experimental animals. To quantify the differences in precision, we define

Table 3.2: Parameter estimates and their standard errors, resulting from the quasi-likelihood fit to the modified bladder cancer data when modeling the treatment effects as piecewise constant with 1 and 2 segments.

	Quasi-Likelihood Analyses: Estimate (Est) and Standard Error (Std.Error)							
	2-Panel Data		4-Panel Data		8-Panel Data		Full Data	
	1-segment							
	Est	Std. Error	Est	Std. Error	Est	Std. Error	Est	Std. Error
β_1	-3.367	0.464	-3.325	0.371	-3.173	0.321	-3.519	0.287
β_2	0.181	0.312	0.186	0.308	0.201	0.308	0.165	0.309
β_3	-0.485	0.281	-0.481	0.267	-0.466	0.271	-0.501	0.274
γ	0.245	0.057	0.245	0.057	0.244	0.057	0.246	0.058
α	0.976	0.132	0.963	0.089	0.915	0.074	1.025	0.062
τ	0.884	0.212	0.884	0.212	0.887	0.213	0.883	0.212
	2-segments: 0-32 months, and 32-64 months							
	Est	Std. Error	Est	Std. Error	Est	Std. Error	Est	Std. Error
β_1^1	-3.139	0.739	-3.300	0.256	-3.134	0.230	-3.588	0.230
β_2^1	0.127	0.339	0.119	0.322	0.127	0.321	0.103	0.320
β_3^1	-0.483	0.308	-0.494	0.294	-0.482	0.294	-0.515	0.293
β_1^2	-2.955	1.282	-3.214	0.488	-2.948	0.442	-3.668	0.436
β_2^2	0.382	0.537	0.387	0.536	0.382	0.536	0.396	0.535
β_3^2	-0.747	0.553	-0.744	0.553	-0.747	0.553	-0.739	0.552
γ	0.283	0.056	0.285	0.057	0.283	0.057	0.288	0.057
α	0.881	0.253	0.934	0.055	0.880	0.030	1.028	0.027
τ	0.887	0.217	0.882	0.214	0.887	0.215	0.876	0.214
δ_2	0.254	0.364	0.253	0.361	0.254	0.362	0.250	0.360
δ_3	-0.615	0.330	-0.619	0.327	-0.615	0.327	-0.627	0.326

Table 3.3: Parameter estimates and their standard errors, resulting from the quasi-likelihood fit to the original bladder cancer data when modeling the treatment effects as piecewise constant with 1 and 2 segments.

Quasi-Likelihood Analyses: Estimate (Est) and Standard Error (Std.Error)

	2-Panel Data		4-Panel Data		8-Panel Data		Full Data	
	1-segment							
	Est	Std. Error	Est	Std. Error	Est	Std. Error	Est	Std. Error
β_1	-2.947	0.477	-3.192	0.356	-3.099	0.299	-3.428	0.275
β_2	0.108	0.301	0.113	0.301	0.111	0.301	0.118	0.302
β_3	-0.551	0.263	-0.545	0.263	-0.547	0.263	-0.540	0.264
γ	0.236	0.056	0.238	0.056	0.237	0.056	0.240	0.057
α	0.880	0.117	0.949	0.077	0.923	0.064	1.015	0.055
τ	0.846	0.193	0.845	0.193	0.845	0.194	0.848	0.193
	2-segments: 0-32 months, and 32-64 months							
	Est	Std. Error	Est	Std. Error	Est	Std. Error	Est	Std. Error
β_1^1	-2.558	0.709	-3.313	0.241	-3.160	0.216	-3.566	0.215
β_2^1	-0.044	0.309	-0.014	0.308	-0.020	0.309	-0.006	0.309
β_3^1	-0.652	0.284	-0.640	0.284	-0.642	0.284	-0.636	0.285
β_1^2	-2.628	1.097	-3.750	0.384	-3.529	0.355	-4.113	0.351
β_2^2	0.699	0.485	0.720	0.484	0.716	0.484	0.728	0.484
β_3^2	-0.443	0.505	-0.435	0.504	-0.436	0.505	-0.432	0.504
γ	0.289	0.055	0.298	0.055	0.296	0.055	0.301	0.055
α	0.747	0.212	0.972	0.043	0.927	0.025	1.047	0.023
τ	0.853	0.197	0.843	0.194	0.844	0.194	0.843	0.194
δ_2	0.328	0.346	0.353	0.343	0.348	0.344	0.361	0.343
δ_3	-0.548	0.311	-0.537	0.309	-0.539	0.309	-0.534	0.308

the design effect (DEFF) for a panel design as the factor by which the sample size must be increased in the panel study so as to achieve equal precision for the estimation of a treatment effect based on the corresponding continuous followup study. Table 3.4 presents the DEFFs corresponding to the six scenarios discussed in Section 2.3 (see Figure 2.3) for 1 and 2 panel designs under the assumption of parameter values close to those obtained in the analysis of the bladder cancer dataset (Weibull parameters $\alpha = 1$ and $\beta_1 = -3.5$; treatment effect $\beta_2 = -0.5$; gender effect $\gamma = 0$; overdispersion parameter $\tau = .9$). For the 2-panel designs that do not achieve joint balance, the DEFF is on the order of 1.3% or less, so a panel study with sample size only 3% larger than M will yield treatment effect estimators with the same precision as from an analysis of continuous followup with M individuals. However, for some designs with current status data (single panel design), design effects up to 50% can be observed. The power curve corresponding to design Bz_b is illustrated in Figure 3.3. Recall that design Bz_b reflects an extreme case of imbalance.

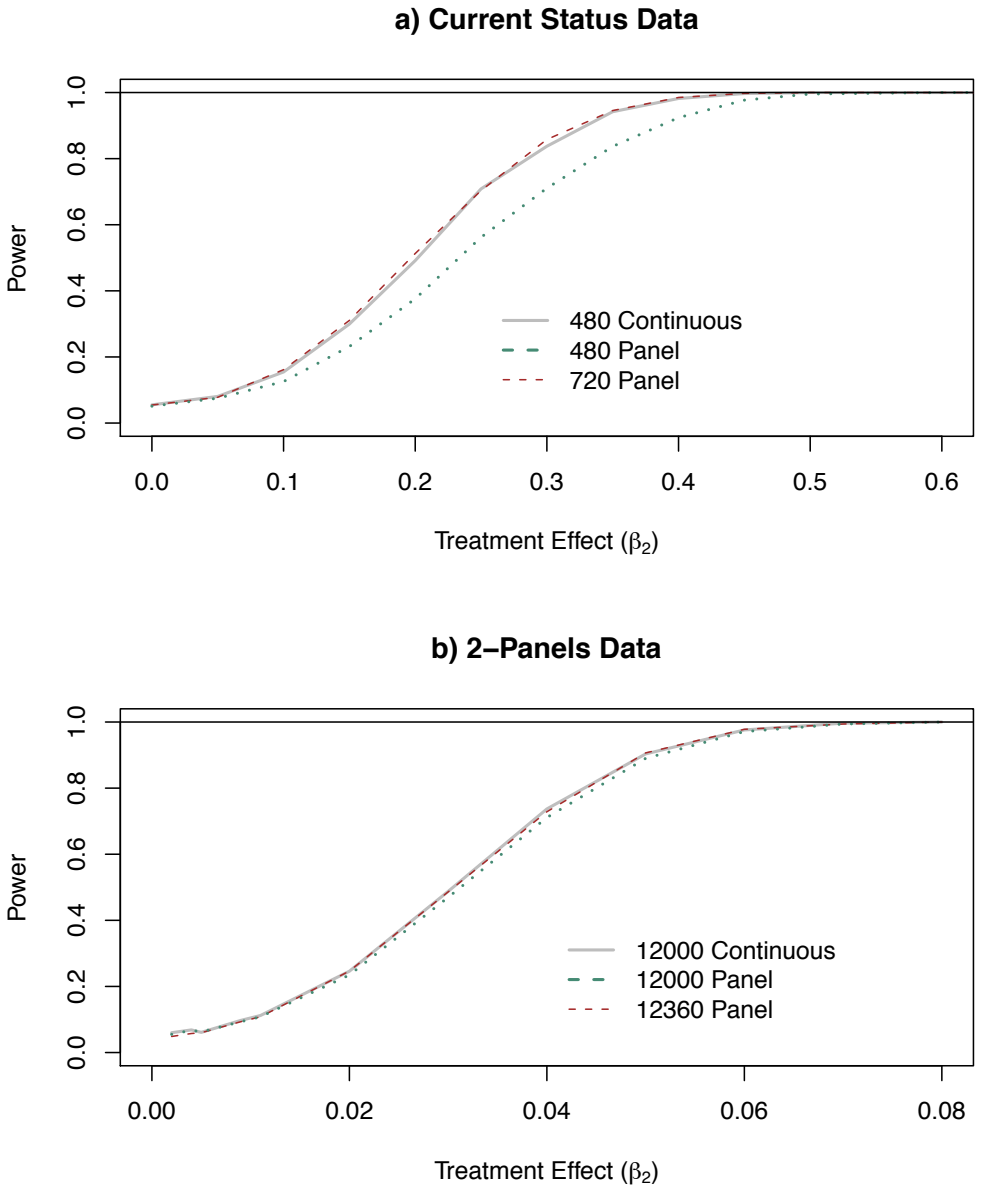
Table 3.4: Design effects for different designs for a similar scenario as the bladder cancer dataset ($\alpha = 1$; $\beta_1 = -3.5$; $\beta_2 = -0.5$; $\gamma = 0$; $\tau = 0.9$) .

Design	Bt*z	Btz	Bt	Bz_a	Bz_b	B
1-panel	1.00	1.01	1.04	1.16	1.50	1.15
2-panels	1.00	1.00	1.00	1.01	1.02	1.01

3.6 Discussion

An alternative use of the methods presented in this thesis relates to optimal assignment and followup of individuals when augmenting a study with new recruits over time. The methods suggest how judicious adjustments to allocation of subjects to treatment groups may help correct for the loss of efficiency due to differential dropout over treatment strata. For example, in a 4-panel study with considerable dropout in one group in the first two panels, the measures of balance presented here could be used to guide the assignment of new individuals so as to achieve higher efficiency for specific parameters.

Figure 3.3: Power of the test that $\beta_2 = 0$ for design Bz_b with parameters $\alpha = 1, \beta_1 = -3.5, \gamma = 0, \tau = 0$. The different curves denote the power for different sample sizes and designs; power with a sample size of 480*1.5 and 12000*1.03 are shown in a) and b), respectively.



Chapter 4

Bayesian Joint Modeling of Zero-inflated Panel Count and Severity Outcomes

4.1 Introduction

Longitudinal count data with marks occur in many scenarios. For example, earthquake data commonly consist of spatio-temporal locations of earthquakes, as well as continuous data on their magnitudes; insurance data may consider recurrences of accidents in a portfolio, as well as insurance costs associated with each event, or with the portfolio. In the medical field, such joint data are less commonly modeled, though count data with marks are certainly not rare. For example, our motivating study considers recurrences of hot flushes following premenopausal ovariectomy and, associated with these, measures of severity of the events. The data arise from a randomized, double-blind trial whose major goal is a comparison of the effectiveness of two treatments in controlling hot flush symptoms.

There has recently been a substantial interest in joint modeling of varieties of discrete and continuous outcomes. In a landmark publication, Dunson (2000) provided a framework for the joint analysis of different types of responses, for example, count and continuous variables. This has been enormously useful, especially for the joint analysis of questionnaire outcomes in sociological and psychological contexts, where several questions in a survey are fundamentally related and measure highly correlated responses. However, this framework is broadly useful in many contexts.

Here, we consider the analysis of panel count data. Panel counts refer to counts accumulated over a particular interval and are the usual outcome recorded in diary studies. In the motivating randomized trial, we collect daily information on whether hot flushes occurred, the number of hot flushes experienced, and whether the experience was rated as high severity, over a period of a year. Patients are randomized to one of two treatment arms. The severity measures for a given day are binary variables indicating a low or high severity experience, and represent an average mark of severity of the hot flushes for the day. The data exhibit several features which need to be accommodated in our analysis. In general, it may be expected that the data are serially correlated with respect to all outcomes measured. Additionally, there is substantial heterogeneity in the outcomes across individuals, and the data are subject to missingness. Importantly, the count data are zero-heavy in both treatment arms, and we expect correlation across the outcomes of counts and severity measures.

In the medical field, when modeling marked point processes, investigators usually focus on one of the two outcomes; for recurrent event processes, counting process methods may be employed, and for modeling the marks, repeated measures approaches are commonly utilized. French and Heagerty (2009) developed GEE methods by assuming that the association between the recurring process and the marks can be explained through an exposure covariate which is measured over time. Cai et al. (2010) proposed an estimating equations approach based on a proportional means model for the marks; they first obtain estimates regarding the recurrent events process, and then substitute these into a model for the marks to estimate the parameters governing that process and to link the processes. With both approaches there is no direct way of testing the association between the two processes. Our approach allows testing for the existence of such association, and consists in a straightforward modeling of both the recurrent event process and the marks through linking random effects.

In Section 4.2, we describe the diary data which motivated this project. In Section 4.3 we develop models for the joint analysis of zero-heavy counts and severities as arising from such diary data. We consider models which use shared frailties to induce correlations across the two outcomes. In Section 4.4 we illustrate our methods through an analysis of the diary data. Section 4.5 describes the results of a simulation study which investigates the bias arising from estimation using separate rather than joint models in situations where

outcomes are linked by shared frailties. The chapter closes with a discussion in Section 4.6.

4.2 Hormone Therapy Study

Counts of the number of hot flushes in a day over a period of one year, along with their severities, were recorded in a clinical trial that compared oestrogen therapy (*conjugated equine oestrogen*: denoted as CEE), the gold standard for the treatment of hot flushes, with *medroxyprogesterone acetate* (referred to as MPA). The participants were healthy menstruating women prior to hysterectomy/ovariectomy for benign disease (Prior et al., 2007) between the ages of 32 and 53 years who were randomly assigned to one of the two treatment groups. There were 20 women in the MPA group and 18 in the CEE group; one of the women from the MPA group was eliminated because of the extreme number of hot flushes reported. The severity random variable reflects the intensity of the hot flush experience in a day, and is an indicator for high intensity. Note that it has been suggested (personal communication, J. Prior) that this variable may reflect the maximum severity associated with each of the hot flushes experienced (instead of an average value) in a day. In addition to the information on counts and severities, we also have the potential prognostic factors: age and BMI of the individual; BMI ranges between 17.69 and 32.62 units.

Figure 4.1 displays the daily data for the number of counts and the severity experience (high/low) over the 364 day period. Red and darker colors correspond to larger counts and high severities in a day. A great deal of heterogeneity is evident in the data. Counts of zeros are represented by gray dots, while gaps indicate missing data. There are fewer events over time for the MPA group, as well as lower severities.

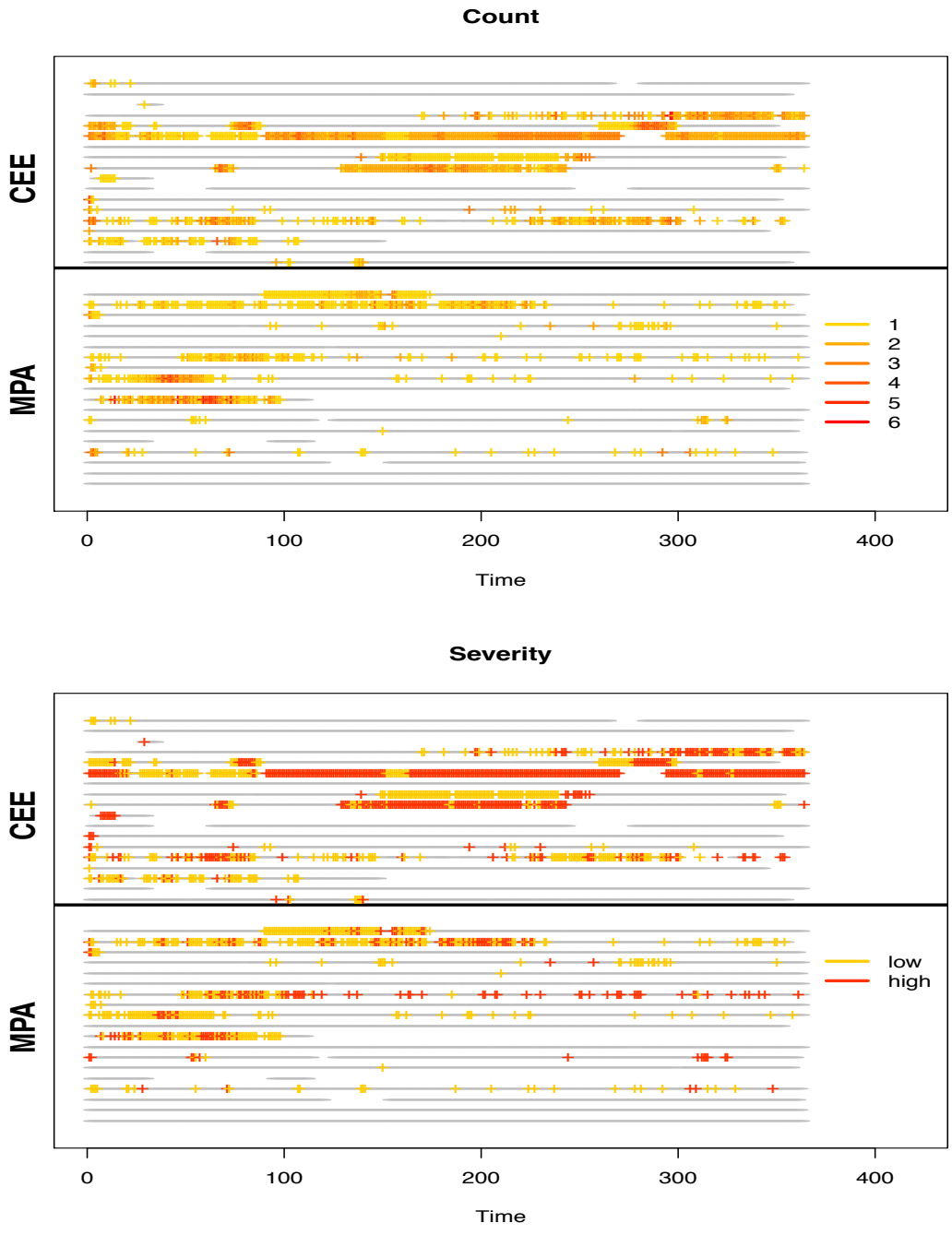


Figure 4.1: Hot flush episodes showing the count level (upper graph) and the severity level per day (lower). A dark shade of red indicates larger counts and high severities. Time is given in days.

4.3 Three-component Joint Model for Zero Heavy Counts and Severities

We describe below a three-component model which distinguishes between the mechanisms that affect the presence of events and the number of events, and relates the severities with the number of events, when events are present. The counts of events are zero-heavy, and we model them using a hurdle model (Ridout et al., 1998; Ainsworth, 2007), which allows flexibility in accounting for a large number of zeros, and, as we will show later, also provides straightforward mechanisms for relating counts to severities.

For individual i, on day d of week t, let Y_{1itd} be an indicator variable for the presence of events, and, if events are present, we also have available responses Y_{2itd}, which records the number of events, as well as Y_{3itd}, an indicator for a high severity day, $i = 1,\ldots,M$, $d = 1,\ldots,7$, and $t = 1,\ldots,52$. We set the model parameters as constant over any day of a week; this provides sufficient flexibility for the analysis conducted here. Hence, conditioning on random, time-dependent, individual-specific frailties $W_{1i}(t), W_{2i}(t)$, and $W_{3i}(t)$ operating on these responses, we specify $\mathrm{E}(Y_{1itd}) = p_{it}$, and given event occurrence, Y_{2itd} is distributed as truncated (at zero) Poisson (μ_{it}), while $\mathrm{E}(Y_{3itd}) = \pi_{it}$. The random effects $W_{ki}(t)$ $k = 1,2,3$, provide the mechanism by which we link the responses, and various forms for these will be described more fully below. Given the random effects, and given $Y_{1itd} > 0$, then Y_{2itd} and Y_{3itd} are assumed independent. Conditional on $W_{1i}(t), W_{2i}(t)$, and $W_{3i}(t)$, and for the case of no missing data, the likelihood kernel simplifies to

$$L = \prod_{i,t}(1-p_{it})^{7-Y_{1it+}} p_{it}^{Y_{1it+}} \times \prod_{i,t}\left[\left(\frac{1}{1-e^{-\mu_{it}}}\right)e^{-\mu_{it}}\right]^{Y_{1it+}}\left[\frac{\mu_{it}^{Y_{2it+}}}{\prod_d Y_{2itd}!}\right] \tag{4.1}$$

$$\times \prod_{i,t}\left[(1-\pi_{it})^{Y_{1it+}-Y_{3it+}}\pi_{it}^{Y_{3it+}}\right], \tag{4.2}$$

where $Y_{kit+} = \sum_{d=1}^{7} Y_{1itd}$, $k = 1,2,3$. For constructing the likelihood note that $Y_{2it+} = \sum_{d=1}^{7} Y_{2itd}Y_{1itd}$ and $Y_{3it+} = \sum_{d=1}^{7} Y_{3itd}Y_{1itd}$. When data are missing at random, the conditional likelihood is obtained in a straightforward manner analogous to the above by omitting the contributions for the terms corresponding to days for which data are missing.

We model the parameters as varying smoothly over time and link them through the

random errors as:

$$\text{logit}(p_{it}) = S_{1i}(t) + W_{1i}(t), \tag{4.3}$$

$$\log(\mu_{it}) = S_{2i}(t) + W_{2i}(t), \text{and} \tag{4.4}$$

$$\text{logit}(\pi_{it}) = S_{3i}(t) + W_{3i}(t), \tag{4.5}$$

where $S_{ki}(t)$ are smoothers which depend on treatment; for example, B-splines are used in the application, with

$$S_{ki}(t) = \sum_{l=1}^{L_k} d_{kl}\phi_{kl}(t) + x_i \times \sum_{l=1}^{L_k} b_{kl}\phi_{kl}(t),$$

where ϕ_{kl} are the B-spline basis functions (see, for example, MacNab and Dean (2001)); d_{kl} and b_{kl} are the spline coefficients, $k = 1,2,3$ and $l = 1,\ldots,L_k$. The treatment indicator x_i is 1 if individual i is in the MPA group and 0 otherwise.

Three forms of relationships are considered for the random error terms, as described below.

1. M_1: *Shared Frailty*: $W_{ki}(t) = u_{kit} + \gamma_k s_{it}$, where $u_{kit} \sim N(0,\sigma_k^2)$, $s_{it} \sim N(0,\sigma_s^2)$, $k = 1,2,3$. Here s_{it} is a shared random effect across all outcomes, and γ_k is the factor loading of this shared effect on outcome k. The u_{kit}'s represent additional heterogeneity beyond the shared random effect. Without loss of generality, we set γ_1 to 1.

2. M_2: *Joint Multivariate Models*: $W_{ki}(t) = u_{kit}$, where $(u_{1it},u_{2it},u_{3it}) \sim MVN(0,\Sigma)$, where all elements of Σ are unknown and estimated.

3. M_3: *Separate models*: $W_{ki}(t) = u_{kit}$, where $u_{kit} \sim N(0,\sigma_k^2)$ independently, $k = 1,2,3$.

Especially when daily data are considered, in addition to accommodating association between responses, the random effects may also accommodate autocorrelation; for example, with $W_{ki}(t) = \rho W_{ki}(t-1) + \varepsilon_{ki}(t)$, where $\varepsilon_{ki}(t) \sim \text{N}(0,\tau_k^2)$, and ρ denotes the autocorrelation parameter.

Implementation of these methods may be carried out in SAS using `Proc NLMIXED`, or through Bayesian MCMC methods (OpenBUGS; Thomas et al. (2006)). Here we utilize

a Bayesian approach, using conventional and independent choices for vague prior distributions and their hyper-parameters. In particular, for all regression parameters we assign $N(0, 10000)$ priors. We assign gamma priors to the inverse of the variance components, $1/\sigma_s^2$ or $1/\sigma_k^2 \sim \Gamma(r,m)$ whose kernel density is $m^r z^{r-1} \exp(-mz)/\gamma(r), z > 0$ with $r = 0.01$ and $m = 0.01$. For the inverse of the covariance matrix, Σ^{-1}, we assign a Wishart prior whose kernel density is $|R|^{p/2}|Z|^{(p-4)/2} \times \exp(-1/2 * \text{Trace}(RZ))$, Z symmetric and positive definite with $p = 3$, and $R = \text{diag}\{1/40, 1 \text{ and } 1/10\}$.

4.4 Analysis of Hormone Therapy Study

The fixed covariates BMI and age were not significant in any of the component models, so we discuss here models without these covariates, thus permitting us to focus on the treatment effect, which was of primary interest. After investigating several forms for the spline components in the models, $S_{ki}(t)$, $k = 1,2,3$, we chose to use, because of their simplicity and parsimony, cubic splines with 4 inner knots at 10, 20, 30, and 40 weeks for p_{it} and π_{it}, and cubic splines with 3 inner knots at 10, 20, and 30 for μ_{it}. Autocorrelation terms were insignificant for all component models and are not discussed here. Finally, we note that models with only individual-specific random effects were also considered, i.e. models for which $u_{kit} = u_{ki}$ and $s_{it} = s_i$, but these provided poorer fits.

Figure 4.2 (left panel) displays the estimated linear predictors for the three component models of M_1: $\text{logit}(p_{it})$, $\log(\mu_{it})$, $\text{logit}(\pi_{it})$ for each treatment arm. The estimates from M_2 are visually indistinguishable from those of M_1 and are not displayed here. The estimated log odds of the presence of events decreases over time for the MPA group, and is generally constant for the CEE arm. When events are observed, their estimated mean for the CEE group seems constant, but decreasing over time for the MPA group, with a clearer difference between these two arms observed from about week 30. The estimated log odds of high severity events is similar for both groups. Figure 4.2 (right panel) also displays the fitted linear predictors from M_3, the separate models. Fits from M_1 and M_3 are quite similar, except for the log odds of high severity. Scatter plots of the relationships between the posterior means of the random effects obtained from M_3 are displayed in Figure 4.3 and indicate a strong positive association between W_{2it} and W_{3it} with other associations being more weakly positive. Notice from the bottom panel of Figure 4.3 that low values of counts tend to be aligned with very low severities.

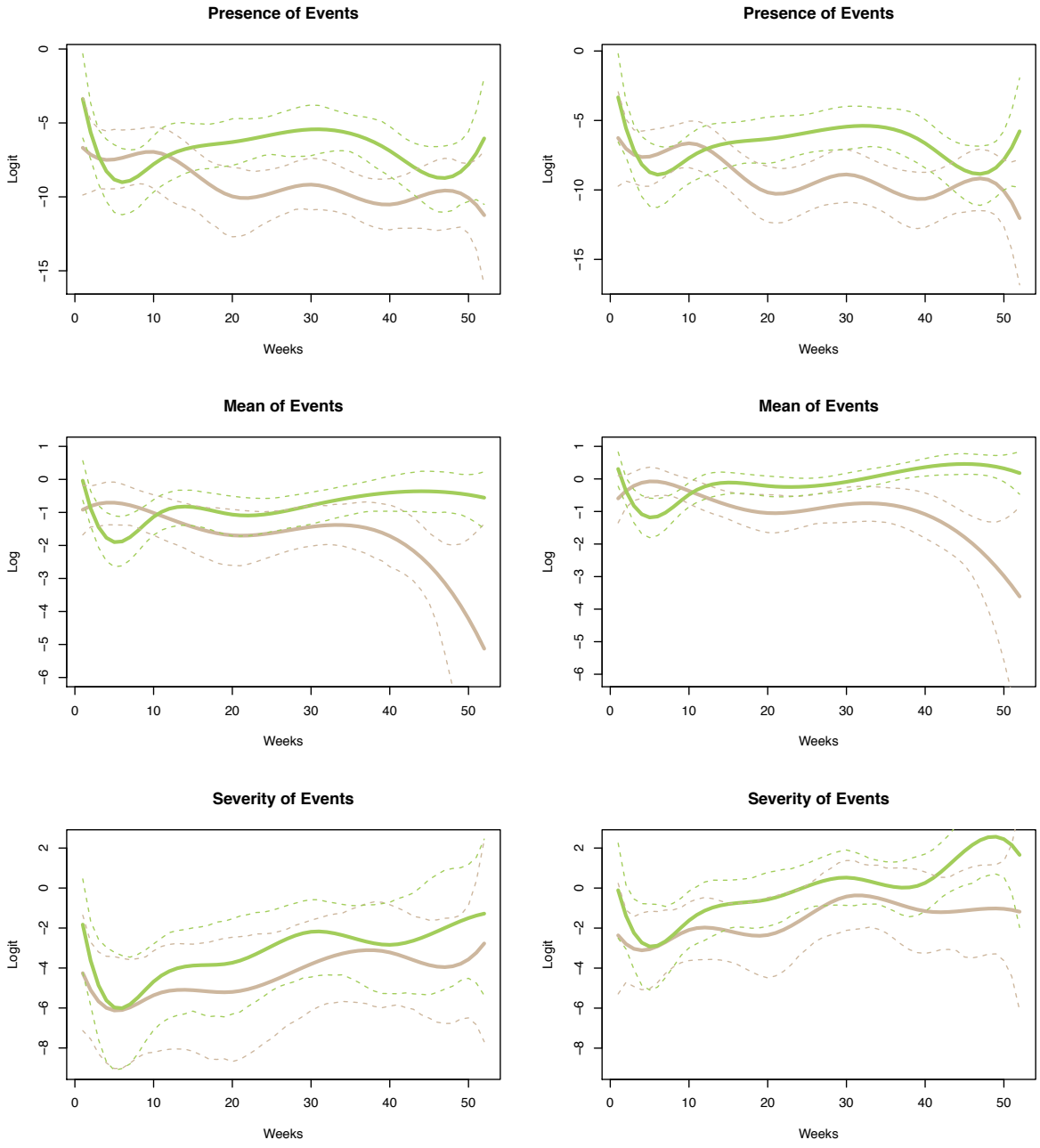


Figure 4.2: Estimated predictors from the fits of models M_1 (left panel) and M_3 (right panel). Rows display estimates of $\text{logit}(p_{it})$ (top), $\log(\mu_{it})$ (middle) and $\text{logit}(\pi_{it})$ (bottom). Estimates for the CEE group/ MPA group are displayed in green/brown solid lines with green/brown dashed lines indicating 95% confidence intervals.

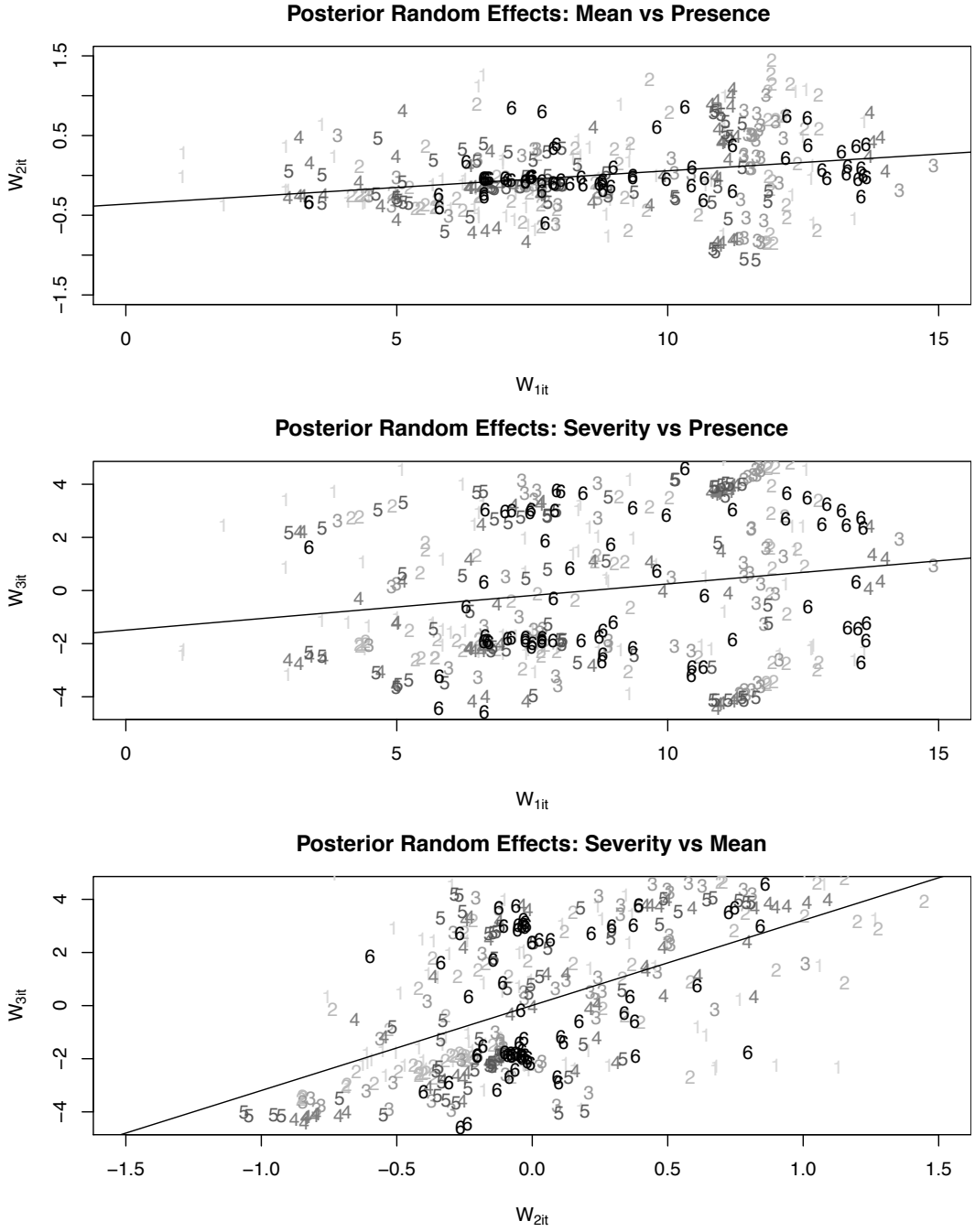


Figure 4.3: Posterior means of the random effects for each individual obtained from the fit of M_3. Only records with events are shown. The numbers in the figures denote the bimester to which the random effect belongs; for instance, '1' refers to the first two months.

Figure 4.4 contrasts the fitted linear predictors from the three models at 4 time points: 12, 24, 36, and 50 weeks. There is general agreement on the direction of the treatment effect for all components of all models, with agreements across the models being closest for estimates of p_{it}. A striking difference between the fitted predictors from the joint and separate analyses is that the log odds of the probability of high severity events is estimated as much lower from both joint models compared to the estimates from M_3. Table 4.1 (panel labelled *HT Study*) lists these estimates at the four time points, as well as the difference between estimates (labelled *Bias?*) obtained from M_1 and M_3. Also listed is the difference between the estimated $\text{logit}(\pi_{it})$ for the CEE group versus the MPA group (labelled *Trt*), which, in contrast, is very similar for the two models.

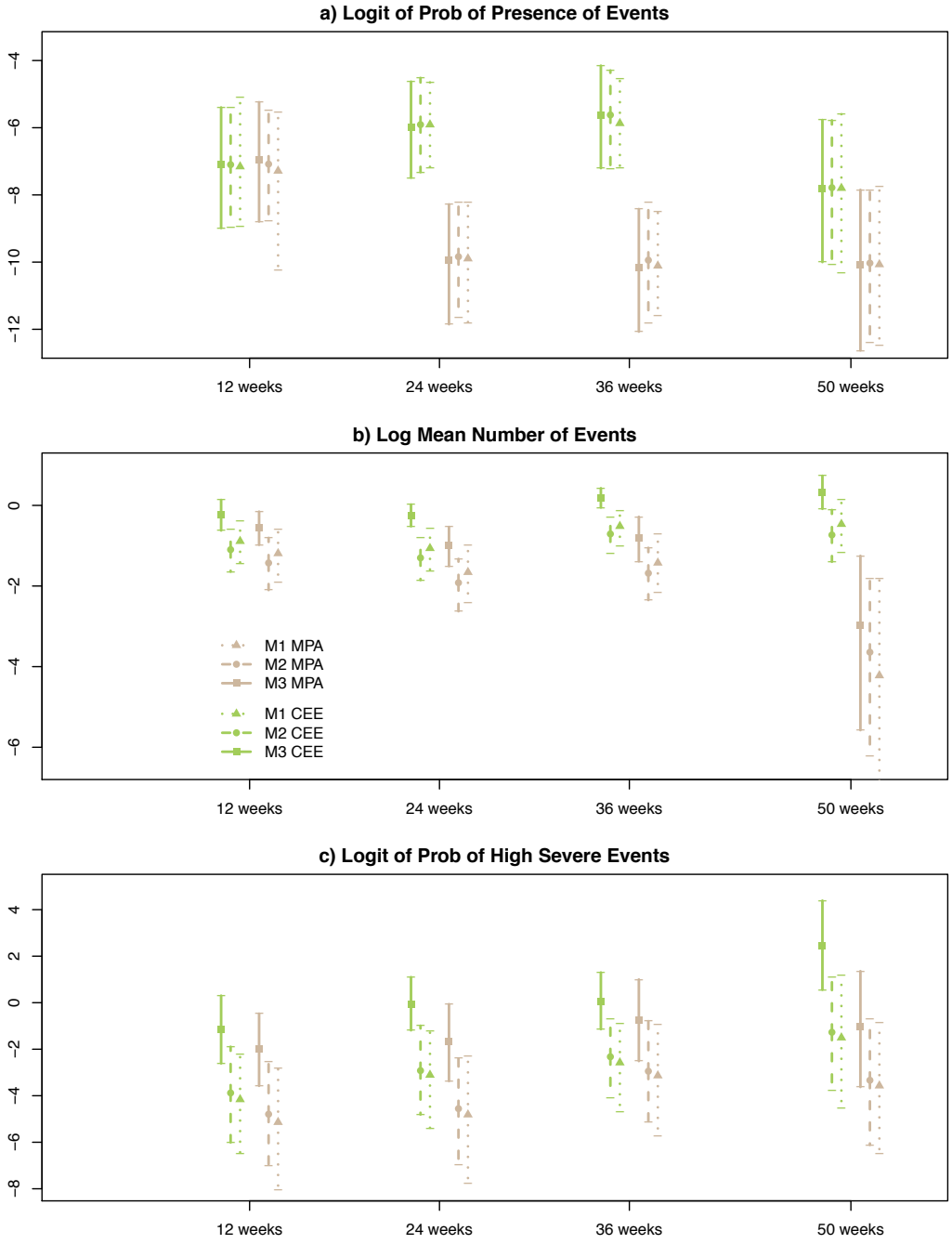


Figure 4.4: Estimated linear predictors at 12, 24, 36, and 50 weeks from M_1, M_2, and M_3.

Table 4.1: Estimated linear predictors of logit(π_{it}) at 12, 24, 36, and 50 weeks from M_1 and M_3 for the hormone therapy analysis, and the simulation investigation. The panel labelled *HT Study* refers to the hormone therapy analysis, while that labelled *Simulated Data* refers to the investigation described in Section 4.5

		HT Study			Simulated Data		
	Week	M_1	M_3	Bias?	M_1 (True)	M_3	Bias
MPA	12	-5.14	-1.97	3.17	-5.17	-2.24	2.93
MPA	24	-4.81	-1.65	3.16	-5.08	-2.16	2.92
MPA	36	-3.14	-0.73	2.41	-3.27	0.05	3.32
MPA	50	-3.57	-1.04	2.53	-3.70	-0.41	3.29
CEE	12	-4.16	-1.13	3.03	-4.27	-1.13	3.14
CEE	24	-3.11	-0.06	3.05	-3.34	-0.35	2.99
CEE	36	-2.58	0.06	2.64	-2.60	0.26	2.86
CEE	50	-1.50	2.44	3.95	-1.79	1.74	3.53
Trt	12	0.98	0.84	-0.14	0.90	1.11	0.21
Trt	24	1.71	1.59	-0.11	1.74	1.81	0.07
Trt	36	0.56	0.79	0.23	0.68	0.21	-0.46
Trt	50	2.07	3.48	1.42	1.91	2.15	0.24

Table 4.2 presents estimates of the variance components and linking parameters for the three models. The estimates of the linking parameters in M_1 are significant providing evidence of association between each pair of models. The correlation between the random effects in the joint shared model may be calculated as

$$\rho_{xy} = \frac{\gamma_x \gamma_y \sigma_s^2}{\sqrt{(\sigma_x^2 + \gamma_x^2 \sigma_s^2)(\sigma_y^2 + \gamma_y^2 \sigma_s^2)}},$$

where $x, y = 1, 2, 3$. These pair-wise correlations, computed from the posterior means, are reported in Table 4.3, and are very close to the ones obtained from the analysis using the joint multivariate model M_2. The three estimated pairwise correlations are all high, but particularly that between Y_{2it} and Y_{3it}, in agreement with the evidence suggested from the fit of M_1 displayed in Figure 4.3.

In Table 4.3, we also report the percentage of the heterogeneity explained by the shared

Table 4.2: Estimates of variance components and linking parameters in all models. The 95% credible interval limits are displayed under CI lower and CI upper.

Parameter	Estimate	Std. Dev.	CI lower	CI upper
		$M1$: Shared Frailty		
γ_2	0.21	0.03	0.16	0.28
γ_3	0.92	0.15	0.68	1.26
σ_1^2	28.87	5.53	19.42	41.04
σ_2^2	0.05	0.04	0.01	0.16
σ_3^2	1.06	0.95	0.01	3.26
σ_s^2	16.69	5.06	7.69	26.98
		M_2: Joint Multivariate		
ρ_{12}	0.67	0.09	0.48	0.82
ρ_{13}	0.54	0.11	0.29	0.72
ρ_{23}	0.89	0.04	0.79	0.94
σ_1^2	44.87	5.68	35.42	57.55
σ_2^2	0.90	0.20	0.58	1.36
σ_3^2	13.12	3.03	7.76	19.70
		$M3$: Separate		
σ_1^2	45.67	5.49	36.16	57.82
σ_2^2	0.51	0.09	0.36	0.72
σ_3^2	11.49	2.24	8.01	16.82

latent variable s_{it} in each component model: about one third in the model for the presence/absence of events p_{it}, and almost all the variability in the other two components of the model. Table 4.3 also presents the Deviance Criterion Information, referred as DIC, (Spiegelhalter et al., 2002) for the different models indicating that both M_1 and M_2 are comparable in their fit of all three model components and provide a better fit than M_3.

Table 4.3: Estimated correlations from joint models and model diagnostics.

Model Type	ρ_{12}	ρ_{13}	ρ_{23}	DIC_{count}	DIC_{sev}	DIC_{total}
M_1	0.59	0.58	0.93	10730	732.4	11462.4
M_2	0.67	0.54	0.89	10730	725.7	11455.7
M_3				10790	748.4	11538.4
% of Latent Variable	0.37	0.94	0.93			

The results for all models were obtained using MCMC methods in OpenBUGS with three parallel chains and a burn-in of 10000 samples; we retained the last 160000 iterations of each chain with thinning set at 40, which yielded a sample of 12000 simulation draws. Graphical monitoring of chains along with their sample autocorrelations were used as diagnostics for convergence.

4.5 Investigation of Bias in Misspecified Shared Frailty Models

The striking difference in estimates obtained from M_3 as seen in the bottom panel of Figure 4.4 and Table 4.1 prompts an investigation of the potential bias in the severity component of the model when a joint model is true, but M_3 is employed. We simulated 1000 datasets from the joint model M_1 with true parameters as the posterior means from the analysis of the hormone therapy study, and with the design exactly as for that study.

Figure 4.5 a) displays the true spline mean for the linear predictor, and the median of the estimates from the simulated datasets. Table 4.1 reports the corresponding information at time points 12, 24, 36, and 50 weeks. The bias is about the same size as that seen in the data analysis (reported also in Table 4.1) and is evident for the overall mean term rather than the treatment effect (the difference between the two splines).

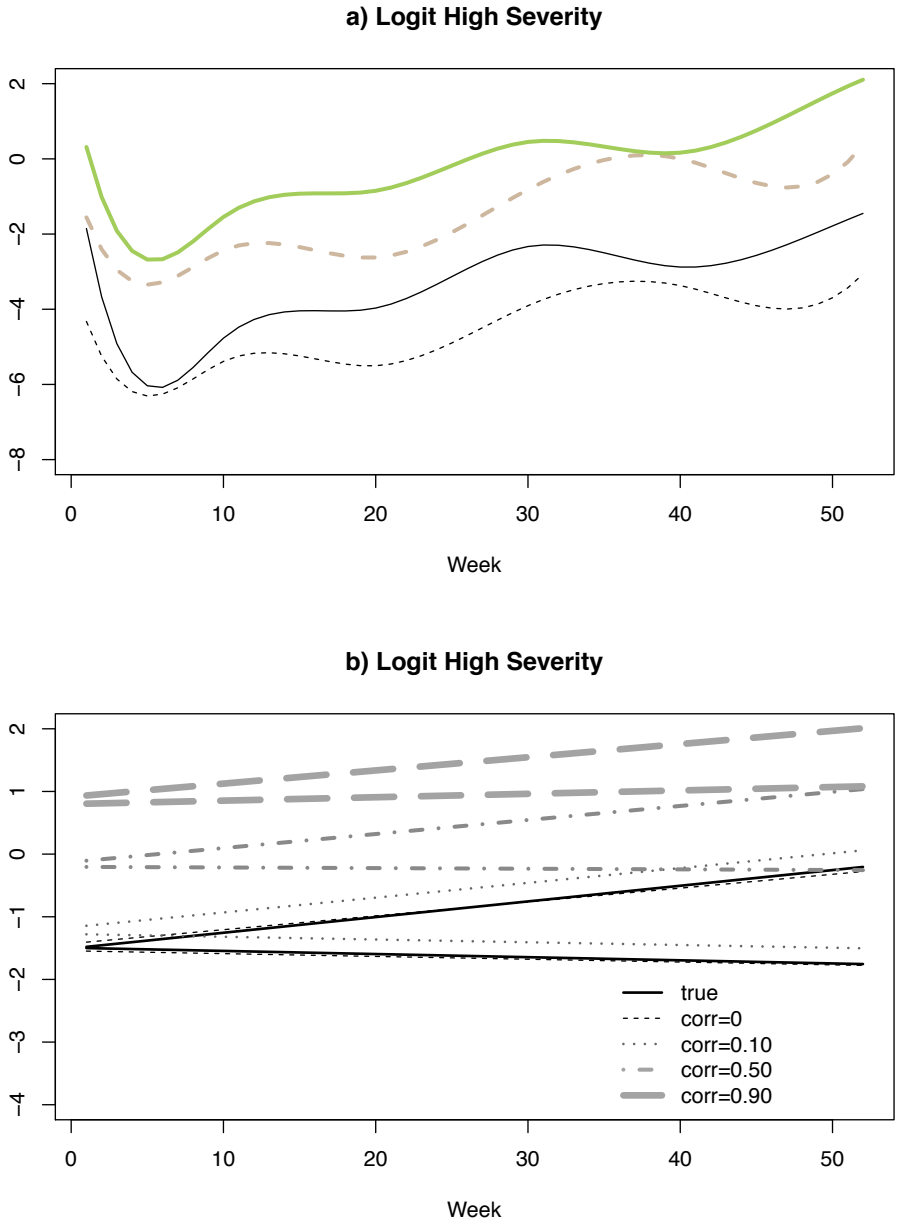


Figure 4.5: a) Linear predictors of logit(π_{it}): black lines correspond to true values, and green (CEE) and brown (MPA) to the median of the estimated values from the fit of M_3 over the 1000 simulations. b) Linear predictors of logit(π_{it}): black lines correspond to true values, and different shades of gray correspond to means of estimates from the fit of the separate model over 250 simulations, where the pair-wise correlations between model components are $0, 0.10, 0.50$, and 0.90.

We also simulated data from a simpler joint model with varying levels of correlation between the random effects in order to assess the effect of the correlation on the bias of an analysis using separate models. For this second investigation we model the predictors as

$$\eta_{kit} = d_{k1} + d_{k2}t + b_{k1}x_i + b_{k2}x_i t + W_{ki}(t) \tag{4.6}$$

where $\eta_{1it} = \text{logit}(p_{it})$, $\eta_{2it} = \log(\mu_{it})$, and $\eta_{3it} = \text{logit}(\pi_{it})$; $x_i = 1$ if individual i is in the treatment group and 0 otherwise. Here again, the sample size and the missing data mechanisms are the same as for the hormone therapy study, and 250 runs were conducted. The marginal variance (including the shared variance) of each of the three components was set to values 44, 1, and 13, for p_{it}, μ_{it}, and π_{it}, respectively, and the values of the linking parameters were set to $\gamma_1 = 1$, $\gamma_2 = 0.15$ and $\gamma_3 = 0.54$; the values of the shared variance component σ_s^2 were specified according to the correlation required with each pair of correlations set to the same values, for simplicity, of 0, 0.10, 0.50, or 0.90. Figure 4.5 b) displays the linear predictor for the true model, and the mean of the 250 estimates of η_{3it} from separate analyses with the four values of the correlation. Table 4.4 presents the mean of the estimates of the linear predictors and the treatment effect at weeks 12, 24, 36, and 50 for all correlation values, as well as the bias. The bias of the predictors increases substantially as the correlation increases, whereas the estimated treatment effect has little bias.

4.6 Summary

This chapter utilizes shared frailty modeling to link outcomes in an analysis of counts and their marks, here termed severities, in a scenario where large numbers of zeros are present. The methods directly enabled an investigation of associations and a determination of which outcomes were most highly linked. In the hormone therapy study, the linked latent variable may well represent the stress experienced by the individual that week; research has shown that hot flushes are heavily associated with stress (Swartzman et al., 1990; Carmody et al., 2006). As well, the methods have prompted an investigation of whether individuals tend to report severe events if they are experiencing many events and whether better training of participants is required in order to monitor these outcomes more effectively.

Table 4.4: Estimates of linear predictors of logit(π_{it}) at 12, 24, 36, and 50 weeks from the fits of separate analyses when the true model includes a shared frailty term with different degrees of correlation between each pair of random effects. Treatment is denoted as *Trt*.

		True	Correlations:				Bias			
	Week	logit(π_{it})	0	0.10	0.50	0.90	0	0.10	0.50	0.90
Trt	12	-1.56	-1.60	-1.33	-0.22	0.86	-0.04	0.23	1.34	2.42
Trt	24	-1.62	-1.65	-1.38	-0.23	0.93	-0.04	0.23	1.39	2.54
Trt	36	-1.68	-1.70	-1.43	-0.24	0.99	-0.03	0.24	1.44	2.67
Trt	50	-1.75	-1.77	-1.49	-0.25	1.07	-0.02	0.25	1.49	2.82
Control	12	-1.21	-1.16	-0.88	0.14	1.17	0.04	0.32	1.35	2.37
Control	24	-0.91	-0.90	-0.60	0.41	1.42	0.01	0.30	1.32	2.32
Control	36	-0.61	-0.63	-0.32	0.68	1.67	-0.03	0.29	1.28	2.28
Control	50	-0.26	-0.32	0.01	0.99	1.97	-0.07	0.27	1.25	2.22
Trt Effect	12	0.35	0.44	0.44	0.36	0.30	0.09	0.09	0.01	-0.05
Trt Effect	24	0.71	0.75	0.78	0.64	0.49	0.04	0.07	-0.07	-0.22
Trt Effect	36	1.07	1.07	1.12	0.92	0.68	0.00	0.05	-0.15	-0.39
Trt Effect	50	1.49	1.45	1.51	1.25	0.90	-0.04	0.02	-0.24	-0.59

Generally, however, these sorts of models may be useful for a variety of settings and it would be important to have a more detailed understanding of the errors which may be incurred through model misspecification in complex analyses such as these. This requires a wide and thorough investigation, both theoretically and by simulation. For practitioners, we recommend simulation confirmation of properties of estimators when conducting related studies using a variety of proposed joint outcome models. It is also conceivable that the shared effects are different for the two treatment arms, and extensions of this nature may also be considered. Further extensions of these type of models for multivariate marked recurrent events may also be considered when recurrences of more than one outcome are of interest. We discuss this in more detail in the context of a criminology application in the next chapter (see Section 5.3).

Chapter 5

Future Work

5.1 Two-stage Designs for Recurrent Events

Several important questions regarding the design of recurrent event studies may be considered as direct extensions of the results provided in this thesis. One important extension would consider adaptive balancing of trials, where methods developed in Chapters 2 and 3 would be utilized to obtain measures of balance in a study, with such measures of balance utilized in deciding strategies for new recruitment in an adaptive, staged, recruitment plan. Additionally, note that there is little discussion in the literature on simple problems as the frequency and timing of followup. Cook (1995) has discussed length of study and sample size based the assumptions of a homogeneous Poisson process with constant accrual rate and exponential independent loss to followup. Matsui and Miyagishi (1999) built upon Cook's (1995) design for an osteoporisis trial where information at the followup times was available only if at least one event had occurred; the rate of occurrence of events was assumed to be piece-wise constant. Matsui (2005) provided sample size calculations for Poisson process models with overdispersion and time-varying treatment effects. Addressing the current challenges of optimizing resources and generalizing assumptions, Cook et al. (2009) proposed a two-stage design to address the issue of not having previous information about expected number of events and the level of heterogeneity between individuals; data from the first stage are used to provide initial estimates of means and heterogeneity to be used for developing sample size calculations for the second stage. Methods developed in this thesis may be used to investigate optimal placement of followup times and the effects

of varying sample size. The focus here is on efficient estimation of the parameters modulating the baseline intensity since these are the ones which are most affected by imbalance. In this regard, we may use flexible semiparametric forms for modeling the intensity function as splines (Nielsen and Dean, 2008). The effect of misspecification of the form of the intensity function could also be explored.

5.2 Assessment of the Use of Scoring Approaches for Analysis of Counts and their Severities

In hot flush studies it is quite common to record the frequency and severity of the symptoms and to evaluate treatments using a single composite measure of both outcomes; this is usually a score constructed as the sum of the number of hot flushes weighted by their severities. Such an approach is also related to analysing joint outcomes as the number of car accidents in a portfolio as well as marks representing the claim amount for each accident through an analysis of the total claim amounts. As well, many other types of health studies have used this type of composite score, with different sorts of ad hoc approaches for weights. See, e.g. Toulis et al. (2009), in which a comparison of the frequency of hot flushes and the composite score was conducted; as well as related work by Nelson et al. (2006), Kaszkin-Bettag et al. (2009), Huang et al. (2006). Because this composite score of frequency and severities of hot flushes is widely used, it would be useful to contrast this approach with the joint analysis and to assess the performance of the use of the composite score in identifying treatment effects. Particularly since the use of the composite score is straightforward and an analysis of such outcomes easy to conduct, researchers working in this field would appreciate guidelines which provide some reassurance of the sorts of situations when such approaches are satisfactory.

5.3 Multivariate Marked Recurrent Events with Exposure

Sometimes, in addition to the presence of marks associated with a point process, there is an exposure variable which needs to be taken into account when describing the occurrence of the events and the development of marks over time. A study which motivates this research is the Pathways to Desistance Study (`http://www.pathwaysstudy.pitt.edu/`); collaborations with the principal investigators of this study have recently been initiated.

In the Pathways to Desistance Study, 1,354 adolescent offenders (between 14 and 18 years old) in two counties of the U.S. were recruited after being found guilty of a serious offense (for example, felonies, sexual abuse, weapons offenses), and were followed for 7 years. The responses are collected at a monthly level on the recurrence of 24 types of crime, as for example, being in a fight, shooting at someone, and robbery, as well as the number of times the crime was committed and a variety of marks associated with the crimes. In addition, there is information on how many days an individual was housed in secured, closed facilities (psychiatric hospital, drug/alcohol treatment unit) where the engagement in crime is less likely. Also available are potential prognostic factors which are classified into the following domains: (1) *indicators of individual functioning*: e.g. work and school status and performance, substance abuse, mental disorder, antisocial behavior; (2) *psychosocial development and attitudes*: e.g., impulse control, susceptibility to peer influence, perceptions of opportunity, perceptions of procedural justice, moral disengagement; (3) *family context*: e.g., household composition, quality of family relationships; (4) *personal relationships*: e.g., quality of romantic relationships and friendships, peer delinquency, contacts with caring adults; (5) *community context*: e.g., neighborhood conditions, personal capital, social ties, and community involvement, and (6) *life-events*: education, income generating activities, living situation, sanctions and interventions, involvement with legal system. The main objectives of this study are identification of clusters of trajectories of recurrence and the characteristics of the different groups of trajectories, as well as how these are explained by social context and developmental changes, and an assessment of the effect of legal intervention (being housed in a secure facility) in promoting changes in trajectories.

It is of particular interest in this study to model and assess the effects of the number of days an individual spent in a secured facility since rates of events are expected to differ over

such periods, especially for specific types of crimes. The study presents several challenges such as the need for methods for joint analysis of a variety of possibly related recurrent event processes, large numbers of zeros, the requirement for the development of methods for clustering recurrent event trajectories, and the presence of marks. Development of models which address these issues would be useful, particularly for determining factors which are associated with crime desistance while adjusting for the exposure to crime. Note that with regard to methods for bi-variate recurrent events, Cook and Lawless (2007) mention three approaches 1) based on time-dependent covariates which may be used to express the dependency between two types of recurrent events 2) based on conditional independence between the types of events, and 3) based on marginal rate or mean models. In this study, there are more than two recurrent event processes to consider, making the setting more complex. We plan to build upon the methods developed in Chapter 4, and explore the use of conditional independence models in which it is assumed there is a latent variable that makes an individual more prone to different events in certain time periods. Further research would also include the development of specific tests for whether the time spent in a secure facility has an effect on the recurrence of crime.

Bibliography

Aalen, O. O., Borgan, O., and Gjessing, H. K. (2008). *Survival and Event History Analysis: A Process Point of View*. Springer, New York.

Ainsworth, L. M. (2007). *Models and Methods for Spatial Data: Detecting Outliers and Handling Zero-Inflated Counts*. PhD thesis, Simon Fraser University.

Andersen, P. K., Borgan, O., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics* **10,** 1100–1120.

Andrews, D. F. and Herzberg, A. M. (2000). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. Springer-Verlag, New York.

Bernardo, J. M. and Smith, A. F. M. (1994). *Bayesian Theory*. John Wiley & Sons, New York.

Byar, D., Blackard, C., and the Veterans Administration Co-operative Urological Research Group (1977). Comparisons of placebo, pyridoxine, and topical thiotepa in preventing recurrence of stage I bladder cancer. *Urology* **10,** 556–561.

Cai, J., Zeng, D., and Pan, W. (2010). Semiparametric proportional means model for marker data contingent on recurrent event. *Lifetime Data Analysis* **16,** 250–270.

Carlin, B. P. and Louis, T. A. (1996). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall, London.

Carmody, J., Crawford, S., and Churchill, L. (2006). A pilot study of mindfulness-based stress reduction for hot flashes. *Menopause* **13,** 760–769.

Cook, R. (1995). The design and analysis of randomized trials with recurrent events. *Statistics in Medicine* **14,** 2081–2098.

Cook, R., Bergeron, P., Boher, J., and Liu, Y. (2009). Two-stage design of clinical trials involving recurrent events. *Statistics in Medicine* **28,** 2617–2638.

Cook, R. and Lawless, J. (2007). *The statistical analysis of recurrent events*. Springer, New York.

Cox, D. R. and Isham, V. (1980). *Point Processes*. Chapman and Hall/CRC, London.

Dabrowska, D. M., Sun, G., and Horowitz, M. M. (1994). Cox regression in a Markov renewal model: An application to the analysis of bone marrow transplant data. *Journal of the American Statistical Association* **89,** 867–877.

Davidian, M. and Carroll, R. J. (1987). Variance function estimation. *Journal of the American Statistical Association* **82,** 1079–1091.

Dean, C. B. and Balshaw, R. (1997). Efficiency lost by analyzing counts rather than event times in poisson and overdispersed poisson regression models. *Journal of the American Statistical Association* **92,** 1387–1398.

Dunson, D. (2000). Bayesian latent variable models for clustered mixed outcomes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **62,** 355–366.

French, B. and Heagerty, P. (2009). Marginal mark regression analysis of recurrent marked point process data. *Biometrics* **65,** 415–422.

Gail, M. H., Santner, T. J., and Brown, C. C. (1980). An analysis of comparative carcinogenesis experiments based on multiple times to tumor. *Biometrics* **36,** 255–266.

Gelman, J. B. Carlin, H. S. S. and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman & Hall, London, 2nd edition.

Huang, M., Nir, Y., Chen, B., Schnyer, R., and Manber, R. (2006). A randomized controlled pilot study of acupuncture for postmenopausal hot flashes: effect on nocturnal hot flashes and sleep quality. *Fertility and Sterility* **86,** 700–710.

Kaszkin-Bettag, M., Ventskovskiy, B., Solskyy, S., Beck, S., Hasper, I., Kravchenko, A., Rettenberger, R., Richardson, A., and Heger, P. (2009). Confirmation of the efficacy of err 731 in perimenopausal women with menopausal symptoms. *Alternative Therapies in Health and Medicine* **15,** 24–34.

Lawless, J. F. (1987). Regression methods for Poisson process data. *Journal of the American Statistical Association* **82,** 808–815.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. Wiley, New Jersey, 2nd edition.

MacNab, Y. and Dean, C. (2001). Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. *Biometrics* **57,** 949–956.

Martinussen, T. and Scheike, T. H. (2006). *Dynamic Regression Models for Survival Data*. Springer, New York.

Matsui, S. (2005). Sample size calculations for comparative clinical trials with over-dispersed Poisson process data. *Statistics in Medicine* **24,** 1339–1356.

Matsui, S. and Miyagishi, H. (1999). Design of clinical trials for recurrent events with periodic monitoring. *Statistics in Medicine* **18,** 3005–3020.

Nelder, J. A. and Lee, Y. (1992). Likelihood, quasi-likelihood and pseudolikelihood: Some comparisons. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **54,** 273–284.

Nelson, H., Vesco, K., Haney, E., Fu, R., Nedrow, A., Miller, J., Nicolaidis, C., Walker, M., and Humphrey, L. (2006). Nonhormonal therapies for menopausal hot flashes. *Journal of the American Medical Association* **295,** 2057–2071.

Nielsen, J. and Dean, C. (2008). Adaptive functional mixed nonhomogeneous Poisson process models for the analysis of recurrent event panel data. *Computational Statistics & Data Analysis* **52,** 3670–3685.

Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*. John Wiley & Sons, New Jersey.

Prior, J., Nielsen, J., Hitchcock, C., Williams, L., Vigna, Y., and Dean, C. (2007). Medroxyprogesterone and conjugated oestrogen are equivalent for hot flushes: a 1-year randomized double-blind trial following premenopausal ovariectomy. *Clinical Science* **112,** 517–525.

Ridout, M., Demetrio, C., and Hinde, J. (1998). Models for count data with many zeros. *Proceedings of the XIXth International Biometric Conference*. *Cape Town* .

Ross, S. M. (1996). *Stochastic Processes*. John Wiley & Sons, New York, 2nd edition.

Spiegelhalter, D., Best, N., Carlin, B., and Van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **64,** 583–639.

Statistics Canada (1996). User handbook and micro-data guide, national longitudinal survey of children and youth. Human Resources Development Canada and Statistics Canada, Ottawa Special Surveys Division, Microdata docu- mentation 89M0015GE, Ottawa, Canada.

Sun, J. and Kalbfleisch, J. D. (1993). The analysis of current status data on point processes. *Journal of the American Statistical Association* **88,** 1449–1454.

Sun, J., Kopciuk, K., and Lu, X. (2008). Polynomial spline estimation of partially linear single-index proportional hazards regression models. *Computational Statistics & Data Analysis* **53,** 176–188.

Sun, Y., Carroll, R., and Li, D. (2009). Semiparametric estimation of fixed effects panel data varying coefficient models. *Advances in Econometrics* **25,** 101–129.

Swartzman, L., Edelberg, R., and Kemmann, E. (1990). Impact of stress on objectively recorded menopausal hot flushes and on flush report bias. *Health Psychology* **9,** 529–545.

Thall, P. F. and Vail, S. C. (1990). Some covariance models for longitudinal count data with overdispersion. *Biometrics* **46,** 657–671.

Thomas, A., O'Hara, B., Ligges, U., and Stutz, S. (2006). Making BUGS open. *R News*. **6,** 12–17.

Toulis, K., Tzellos, T., Kouvelas, D., and Goulis, D. (2009). Gabapentin for the treatment of hot flashes in women with natural or tamoxifen-induced menopause: a systematic review and meta-analysis. *Clinical Therapeutics* **31,** 221–235.

U.S. Department of Labor, Bureau of Labor Statistics (2005). National longitudinal surveys handbook. National Longitudinal Survey of Labor Statistics. Available at `www.bls.gov/nls/handbook/nlshndbk.htm`.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss–Newton methods. *Biometrika* **61,** 439–447.