STUDY DESIGN AND METHODS FOR EVALUATING SUSTAINED UNRESPONSIVENESS TO PEANUT SUBLINGUAL IMMUNOTHERAPY

Monica Chaudhari

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Public Health in the Department of Biostatistics in the Gillings School of Global Public Health.

Chapel Hill
2017

Approved by:

Michael R. Kosorok

Gary Koch

Michael G. Hudgens

Donglin Zeng

Edwin H. Kim

ABSTRACT

Monica Chaudhari : Study Design and Methods for Evaluating Sustained Unresponsiveness to Peanut Sublingual Immunotherapy
(Under the direction of Michael R. Kosorok)

The length of time off-therapy that would represent clinically meaningful sustained unresponsiveness (SU) to peanut allergen remains undefined. Our work has three-fold objectives: first, to delineate aspects of the altered clinical trial design that would allow us to assess effectiveness of sublingual immunotherapy (SLIT) in achieving SU; second, to discuss methodology for evaluating the time to loss of SU and associated risk factors in context of the proposed study design; finally, to develop a flexible methodology for assessing mean reverting threshold and prognosis of SU failure in the presence of study risk factors. Failure refers to the loss of SU upon therapy cessation in peanut allergic children who are administered sublingual immunotherapy (SLIT).

The salient feature of the new design is the allocation scheme of study subjects to staggered sampling timepoints following therapy suspension when a subsequent food challenge is administered. Due to a fixed sequence of increasing allergen doses administered in a challenge-test, the subjects true threshold at either occasion is interval-censored. Additionally, due to the timing of subsequent DBPCFC, the time to loss of SU for subjects who pass the DBPCFC at study entry is either left- or right-censored. In this thesis, we elaborate on the features of the study design, develop and extensively validate methods to evaluate study end points and discuss their potential to inform individualized treatments.

The thesis is compartmentalized as follows: (i) an innovative clinical trial design that aims at studying SU to SLIT; (ii) a newly developed mixture proportional hazards model for evaluating the time to loss of SU in context of the study generated interval-censored data subject to instantaneous failures; (iii) a time-dependent Ornstein Uhlenbeck (OU) diffusion process for modeling immunologic

SU degradation trajectories using stochastic differential mixed effect model (SDMEM) framework; (iv) the estimation of mean-reverting threshold and prognosis of the loss of SU; (v) lastly, the clinical implementation and future scope of work. Through this work, we are presented with an opportunity to dedicate these inter-connected parts to three core issues of failure: model description, prediction and prevention.

To my husband, Amol, and my daughter, Ruhi Ariyah
and
In loving memory of my Grand Parents, Bai and Jeesahib

ACKNOWLEDGEMENTS

It is a pleasure to thank many people who made this thesis possible. First and foremost, I would like to take this opportunity to express my sincere appreciation and deepest gratitude to Dr. Michael R. Kosorok, my advisor, for his inspired guidance, patience, encouragement throughout the ups and downs of this work, and generous help, both academic and non-academic, that made the completion of this dissertation possible. The experience I have gained working with him is invaluable.

I gratefully acknowledge the advice and helpful suggestions of my committee members and thank them for their numerous, enriched ideas pertaining to this work. In particular, I extend my thanks to Dr. Edwin H. Kim who untiringly and enthusiastically took time in educating me the clinical aspects of the problem and helping contextualize the proposed methodology.

Finally, I thank my parents for their continuous encouragement, support, trust, and unconditional love even from a great distance. And, of course, I thank my husband for his support and my little daughter, Ruhi, without whose cooperation this work would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SU	Sustained Unresponsiveness
SDMEM	Stochastic Differential Mixed Effect Model
SLIT	Sublingual Immunotherapy
DBPCFC	Double Blind Placebo Control Food Challenge
MCRT	Minimal Clinically Relevant Threshold
NOAEL	No Adverse Effect Dose Level
LOAEL	Lowest Observed Adverse Effect Dose Level
OU	Ornstein Uhlenbeck
FPT	First Passage Time

CHAPTER 1: INTRODUCTION

Allergic responses to peanut are among the most prevalent, deleterious and long lasting of all food allergies [8; 91]. Prevalence has been rising over the past decade, particularly in the United States pediatric population [72]. The IgE-mediated allergic response is characterized by acute onset of symptoms generally within 2 hours of peanut exposure and can progress to more serious symptomatic sequelae such as anaphylaxis, hypotension, and multiple organ dysfunction syndrome [8; 91]. While the current standard of care for peanut allergy is a strict avoidance of peanut and ready access to self-injectable epinephrine, implementation of an avoidance diet is complex leaving majority of patients with a life-long risk of allergic reactions and even death due to unintentional ingestions [33]. Therapeutic interventions that provide everlasting defense against accidental peanut ingestion are thus needed [76].

Lately, the use of allergen immunotherapy (AIT) is gaining grounds in clinical research for the treatment of food allergy. Based on current research, it is unclear whether permanent tolerance and cure can be be achieved. As a result, the objective, in general, is to increase the reaction threshold in food-allergic individuals to prevent a catastrophic response following inadvertent exposure. Although AIT has long been used as treatment for aeroallergens and hymenoptera venoms, no licensed immunotherapy products are available for the treatment of food allergy. Investigators are pursuing several different routes of its administration including subcutaneous, oral, sublingual, and epicutaneous to treat individuals presented with allergic symptoms to IgE mediated food (6). In contrast to the subcutaneous immunotherapy that has evidenced an unacceptably high rate of systemic reactions, oral immunotherapy has shown some promise but needs more investigation due to more common oral and gastrointestinal disease effects. On the other hand, sublingual immunotherapy (SLIT), which involves the administration of small amounts (micrograms to milligrams) of allergen extract under the tongue and is the locus of our investigation, has been claimed safer with fewer systemic reactions and easier mode of administration.

The Double-Blind Placebo-Control Food Challenge (DBPCFC) test is considered the gold standard in diagnosis of food allergy. The test exposes a patient to a fixed sequence of increasing dose levels of an allergen until he/she demonstrates clinical symptoms. The lowest amount of the allergen that evokes objective symptoms is termed as the eliciting dose (ED). One approach to evaluate the effectiveness of AIT in an investigation is by administering repeated DBPCFCs, one before and one after the course of AIT, to assess the change in threshold of allergic sensitivity (ED) from baseline. Typically, the primary endpoint in these studies is the degree of change in threshold in the treatment group compared with the placebo group. Specific ED based measures of clinical efficacy of AIT include more commonly studied desensitization, which refers to an increase in reaction threshold while receiving the study drug. Tolerance, on the other hand, refers to a non-reactive state of the immune system that persists indefinitely after discontinuation of the study drug. Even though clinical desensitization is demonstrated in almost all OIT and SLIT studies for food allergy, tolerance has less commonly been tested and requires further study. The problem lies more in that there are no well-defined markers proving tolerance; the DBPCFC off of therapy is the current best attempt at assessing tolerance. A recent study [11] suggests that what we have previously thought to represent immunological tolerance may be transient and hence, the term sustained unresponsiveness (SU) is coined to refer to the phenomenon of a non-reactive state that persists after discontinuation of therapy but wanes after a period. In other words, SU is the capacity to maintain desensitization to the food allergen following cessation of therapy. With scarce published literature defining and characterizing SU, the length of time off-therapy that would represent clinically meaningful benefit remains undefined. Appropriate clinical efficacy parameters and study endpoints to demonstrate SU have, therefore, not been established. With the FDA approval of food immunotherapy potentially on the horizon, and an extensive longitudinal data that have been collected during an ongoing randomized clinical investigation of pediatric population for peanut allergy conducted in North Carolina, we are presented with a unique opportunity to address issues pertaining to SU with an innovative study design and advanced statistical methodology.

In this dissertation, we extend the ongoing investigation of the desensitized pediatric cohort towards two-fold objectives: first, to propose and discuss a unique clinical trial design that would allow us to study SU; second, to explore and develop methodology to evaluate the time to loss of SU and the mean reverting threshold in context of the data originating from the proposed study design.

We propose two new statistical procedures that exploit data characteristics emanating from a typical AIT trial and apply those methods to improve models for estimation and prognosis of SU in AIT studies.

The rest of this dissertation is organized as follows. In Chapter 2, we provide a brief introduction to the study design and motivation to develop methods in context of the data arising from this clinical trial. It is then followed by a summary of the literature review of the existing approaches to solving problems like the one described above. This includes an introduction of a mixture model and a computationally efficient algorithm that can be used to analyze current-status data subject to instantaneous failures, and finally, a brief discussion of stochastic differential models with mixed effects with focus on time-dependent Ornstein-Uhlenbeck Process to model individuals SU degradation trajectories. In Chapter 3, we provide a comprehensive methodology that can be utilized to analyze interval-censored data subject to instantaneous failures; in particular, we examine a recently proposed mixture model and a computationally efficient algorithm developed for fitting the proposed model. In Chapter 4, we provide an in-depth discussion of the proposed study design and investigate the finite-sample performance and power of the procedure proposed in Chapter 3. Chapter 5 develops method to model SU degradation trajectories of individuals using Stochastic Differential Mixed Effect Model (SDMEM) framework with a focus on Ornstein-Uhlenbeck latent process, implements estimation using Metropolis-Hastings algorithm and ends with discussing future directions.

CHAPTER 2: BACKGROUND

To provide motivation for this work, we start with a brief discussion of clinical trial design, focusing on the alterations and the rationale that inspired the revised augmentation to the original trial design.

The remainder of this chapter is organized as follows: in Section 2.2, we discuss peculiarities of the data expected to originate from the ongoing clinical trial. This includes discussion of different types of interval-censored data and other key aspects of the study relevant for the development of proposed methodology. Section 2.3 reviews existing approaches for the regression analysis of the interval-censored data and introduces recently developed mixture PH model for the analysis of current-status data with instantaneous failures. This is then followed by a review of SU degradation modeling using time-dependent Ornstein-Uhlenbeck Process under the framework of stochastic differential mixed effect models.

2.1 Study Design

The Peanut SLIT and TLC Study is a revised augmentation to a pre-existing phase-II, prospective, randomized, open label 66-month clinical study supported by the National Institutes of Allergy and Infectious Diseases (NIAID) and conducted in the Allergy, Immunology and Rheumatology (AIR) Division in the Department of Pediatrics at the University of North Carolina at Chapel Hill [26]. Protocols were approved by the Human Subjects Committee at the University of North Carolina. We note that opposite of most studies that focus on estimating treatment effect, the study has an objective to assess the effect of taking away therapy. Following the 48-month DBPCFC, the original protocol mandated that the desensitized subjects be randomized in a blinded fashion 2:1 to placebo or continued peanut SLIT for six months. To assess clinical SU, the placebo group would complete the trial with 54-month DBPCFC, and the treatment group would discontinue open label SLIT for an additional six months to complete the trial with 66-month DBPCFC. However, based on the results

from [83] that showed only 50% SU after a month of using the presumed more robust OIT treatment lead to suspicion that six months off SLIT would not work. Recognizing that the six-month period to assess clinical tolerance was rather optimistic, arbitrary and less supported by prior research, and possibility of negligible success rate would then translate into unacceptable risk and ethical concerns, the protocol was revised with staggered sampling times beyond the dose maintenance period for administration of final DBPCFC among desensitized subjects. This dissolved the need for a two-arm trial and shortened the length of the overall clinical trial. Since time to loss of SU is an important dimension that, in our knowledge, no prior study has investigated, the revised study offers a robust framework and a systematic approach to study this clinically valuable information. In summary, a randomized clinical trial is being conducted to evaluate the safety and efficacy of peanut SLIT in inducing clinically SU among peanut allergic children.

Briefly, the study consists of a screening visit, baseline visit (sometimes combined with the screening visit), build-up phase (approximately, 20 weeks), maintenance phase (42 months) and lastly, SU phase (17 weeks). Upon enrollment, children between age 1 to 11 years underwent an entry DBPCFC with 1000 mg of peanut protein to confirm the peanut allergy diagnosis and establish a baseline threshold level. Following a positive DBPCFC, each subject was required to begin peanut SLIT at a starting dose of 1 pump of a 1/100 dilution of peanut concentrate (~ 1.5 mcg peanut protein). During the build-up phase which lasted approximately 20 weeks, subjects were dosed daily and administered increased number of pumps every 1-2 weeks as per the dosing schedule. Subjects returned to the research unit for observed dosing with each change in peanut SLIT dilution (1/100, 1/10, full concentration) and with every other dose increase on full concentration until the maintenance dose of 16 pumps of full concentration peanut SLIT (4000 mcg peanut protein) was achieved. During the maintenance phase which is underway, subjects continue daily administration of the maintenance dose to return every 6 months for follow-up. At the end of this phase i.e. after at least 48 months of peanut SLIT, subjects are required to undergo a second DBPCFC to 5000 mg of peanut protein to assess desensitization. The desensitization criterion used for qualifying subjects for the following phase of this study include meeting the minimal clinically relevant threshold (MCRT) of 300 mg of peanut protein without symptoms [5]. Those who do not qualify are required to stop peanut SLIT and conclude the study, with a recommendation to resume a strict peanut avoidance diet. The desensitized subjects enter the SU phase during which peanut SLIT is discontinued. They

are block randomized to one of the 17 weeks when the third and final DBPCFC, similar in structure to the 48-month DBPCFC, is administered. The purpose of the final DBPCFC is to assess for SU. This marks study completion and at the primary investigators clinical discretion, subjects are recommended to transition to a daily peanut food equivalent to maintain the desensitized effect.

Due to the protocol revisions mid-way through the study, 24 of the total 51 who were awaiting 48-month DBPCFC, had an extended maintenance phase for a maximum of 6 months beyond the initial 42 months. For analyses, this extension is disregarded based on prior evidence that suggests negligible benefit of continued dosing on the overall treatment outcome.

2.1.1 Double Blind Placebo Controlled Food Challenge (DBPCFC)

Each study subject is administered the protocol DBPCFC at baseline, at the end of 48th month and once during the 17-week SU phase. A nurse or physician who administers DBPCFC is blinded to the testing material and so is the supervising investigator. Before each challenge, the subject is required to have a physical exam and peak expiratory flow measurements performed.

The DBPCFC consists of randomly ordered administration of two parts: one, that consists of graded doses of peanut flour and the other, of identical graded doses of placebo in the form of oat flour. The doses are given every 10-20 minutes up to a cumulative dose of 1000 mg (25 mg, 50 mg, 100 mg, 250 mg, 575 mg) during the entry challenge and up to a cumulative dose of 5000 mg (100 mg, 200 mg, 500 mg, 800 mg, 1300 mg, 2100 mg) at the 48-month and final DBPCFCs. A minimum 10-minute observation period is allowed between doses to monitor for symptoms. Notably, unlike latter challenges, the entry challenge is administered to a lower dose simply to prove allergy status and not the treatment response.

Before resuming the second part of the challenge the subject is observed for a minimum period of 1 hour to allow for washout of dosing and the side effects of the first part of the challenge. Reactions are scored using a Food Challenge Symptom Score (FCSS) sheet. If the subject begins to have significant objective or persistent subjective symptoms, the food challenge is terminated and the subject is given appropriate treatment. Subjects who are symptomatic and receive treatment are observed for a minimum of 2 hours after the challenges are completed before being discharged from the clinical research unit.

2.1.2 Randomized Weekly Assignment for the Final Food Challenge

Before the start of the SU phase of the study, a computer-generated block randomization schedule was prepared such that each subject had an equal probability of being randomized to one of the following time periods: initial 6 weeks, between 7-12 weeks and between 13-17 weeks. A subject was then assigned to a week sampled from the block of weeks they were randomized to. Based on this randomization schedule, three desensitized subjects were randomly assigned to one of the final DBPCFC administered each week during the 17-week SU phase. No desensitized subject is administered a final challenge dose higher than the highest tolerated dose in the 48th month DBPCFC for the ethical purposes of ensuring safety and because in this age group, continued increase in threshold without treatment would not be expected.

2.2 Data Peculiarities: Case of Interval Censoring

This analysis is complicated by interval censoring due to the way the food challenges are administered. In contrast to the missing data that provides no information, censoring refers to an observation in the study being incomplete and hence, provide some information; that is, the quantity of interest is observed only to fall into a certain range instead of being exactly known [38]. There are three different types of censoring: left, right, and interval. Left censoring, usually a rare encounter, occurs when the event has already occurred before the earliest study time. Right censoring, the most common of all, occurs when the event does not occur before the study ends. This usually happens due to time constraints or resource limitations. Interval censoring occurs when the exact quantity is unknown but is known to occur within some range. Right and left-censored data are specific types of interval-censored data. However, the statistical methods for right-censored data do not apply to the interval-censored data and in general, interval-censored data is more challenging to compute.

This study presents a case of triple interval censored data: first, subject's true threshold is obscured by the nature of food challenges, it is bounded by no observed adverse event dose level (NOAEL) and the lowest observed adverse event dose level (LOAEL). In other words, NOAEL is defined to be the highest administered dose at which no adverse event is recorded and LOAEL is the lowest administered dose at which an adverse event is recorded [77]. Therefore, subject's true threshold is interval censored and could be any value within the interval defined by NOAEL and

LOAEL; second, in the absence of therapy, since it is assumed to be guided by a monotonically non-increasing latent process, the time to loss of SU to a targeted dose administered at staggered observation times is either left- or right-censored and depends on rate at which the individual's threshold is likely to fall once he/she is off-therapy. Interval-censored data is classified as either Case-1 or Case-2 data.

2.2.1 Case-1 Interval-Censoring with Instantaneous Events

Case-1, also known as current-status data, usually refers to the case when the only knowledge that exists on the quantity of interest is whether it is lower than the observed quantity [36]. In context of our study design, the only known knowledge of the time to loss of SU is whether it has occurred before the observation time when the final DBPCFC is administered. Unlike right censoring, the exact value of the survival time is never observed. In other words, if $T \in (L, R]$ where T is the survival time of interest, then for Case-1 censored data, either $L = 0$ or $R = \infty$.

2.2.2 Case-2 or Case-k Interval-Censoring

Case-2 interval-censored data includes at least one interval in which both endpoints belong to $(0, \infty)$ [38; 36]. Thus, we only know that T has occurred either within some random time interval, or before the left end point, or after the right end point of the interval. More precisely, with the two observed endpoints, L and R, the data observed is: $(I_{D \leq L}, I_{L<D \leq R}, L, R)$. In context of our study, the threshold data of food challenge is Case-2 interval-censored, where D represents the unobserved true dose threshold. Case-1 interval-censored data is a special case of Case-2 interval-censored data when either $L = 0$ or $R = \infty$. Case-k interval censoring arises when there are k interval-censored observations per subject. This is a generalization of case-2 interval censoring. Our study does not present this kind of censored data.

2.2.3 Independence Assumption

For interval-censored data, like the right-censored data, there is an independence assumption. It is assumed that the mechanism that generates the censoring is independent about the event time of interest [38]. By virtue of the study design, which has the final DBPCFCs at staggered sampling examination times, the observational process is independent of the failure times, which is required by

most approaches to the analysis of interval-censored data. Essentially this says that, except for the fact the T lies between L and R, the interval does not provide any extra information about T.

2.2.4 Instantaneous Failures

At the 48-month DBPCFC, subjects who exhibit threshold lower than MCRT or the targeted dose level are assumed to have lost tolerance for the targeted dose at the study initiation. This could be construed as instantaneous failures suggesting that some subjects can never reach the targeted threshold. The occurrence of instantaneous failures leads to some probability mass at time zero for the distribution of time to loss of SU.

2.3 Methodology: Existing, New and Future Work

In this section, we develop methodologies to analyze two primary endpoints: the time to loss of SU to a given target dose threshold and the mean reverting threshold, with focus on assessing their association with risk factors based on data arising from the clinical trial, and to estimate the baseline survival in the absence of covariate effect. It is worthwhile to note that these objectives are studied through events during the SU phase of the clinical trial. Therefore, for analyses, we mark the 48-month DBPCFC as the "$zero$" time-point embarking the study. We start with a straightforward but a more constrained approach of defining our endpoint. By using more confined information than that at our disposal, we devise a technique that serves an approximate solution. We later build more sophisticate techniques capitalizing on threshold distribution from the challenge tests to find solution to relatively more relaxed problem.

2.3.1 Mixture Proportional Hazards Model

For this method, we confine the scope of efficacy endpoint, time to loss of SU, to only those dose levels that are administered during the final DBPCFC. Further, we assume that subject's true threshold upon cessation of therapy is guided by a monotonically non-increasing latent process. Then, as discussed in the earlier section, the endpoints are not observable but rather are known relative to the staggered sampling time-points of the final DBPCFC, resulting in $case-1$ interval censored data, with non-informative censoring times. Additionally, subjects who exhibit a threshold lower than MCRT

or the targeted dose, are assumed to have lost tolerance for the targeted dose at the study initiation; i.e., in the context of evaluating time to loss of SU, these subjects experience what is commonly referred to as an instantaneous failure. To analyze the resulting $case - 1$ interval-censored data with instantaneous failures, we examine recently developed mixture model, which is a generalization of the semi-parametric PH model. Under the standard PH assumption, the model assumes that the covariates provide for a multiplicative effect on the baseline risk of experiencing a failure both at study initiation and thereafter.

Originally proposed by [17], the proportional hazards (PH) model is (arguably) the most widely used method for the purposes of analyzing time-to-event data. Unfortunately, the underpinnings of the semi-parametric variants of this model to allow for the analysis of interval-censored (IC) data has proven to be more challenging since the time of event is always obscured. Several advances have been made to analyze interval-censored data under the PH model. Methods include joint estimation of regression parameters and the baseline hazard function using a Newton-Raphson based algorithm [25], a marginal likelihood approach [69] and Monte Carlo expectation maximization (EM) algorithm [28] for the estimation of regression coefficients alone, Non-parametric maximum likelihood methods such as Turnbull's self-consistency algorithm [80], iterative convex minorant (ICM) algorithm [32], generalization of Rosen algorithm [66; 93] that primarily focus on estimation of baseline survival or the failure time distribution function in the absence of covariate effect. Later, techniques such as ICM reformulated generalized gradient projection method [60] and a semi-parametric alternative based on multiple imputation [61] allowed for the estimation of covariate effects. More recent approaches include modeling formulations of baseline hazard allowing for covariate effects using expectation-maximization (EM) algorithm [27], local likelihood techniques [7], penalized likelihood with a piecewise-linear spline [14], monotone B-splines [92]. Methods that focus on current status data pertaining to a single failure time of interest include proportional hazards model of [34], proportional odds model using sieve method [35], a computationally efficient EM algorithm [53; 87]. A compendium of methods relating to analysis of IC data could be found in [38; 36; 94; 47].

The bulk of the aforementioned work can either be computationally resource intensive or have difficult implementation. As a result, analysts tend to adopt well established partial likelihood method (*coxph* in R, *phreg* in SAS) used for right-censored data by using the midpoint or the

right endpoint of the observed interval as the exact failure time for the left- and interval-censored observations [31; 4]. However, such an approach has been demonstrated to result in biased estimation and inference [67; 58]. Further, none of these techniques are designed to account for the effect of instantaneous failures. The proposed analyses is built upon a more flexible, robust to initialization and computationally efficient semi-parametric methods of Wang et al, 2016 [87] with extensions detailed in recent Withana Gamage et al, [89], that account for the special case of $case - 1$ interval-censored data subject to instantaneous failures. For the ease of implementation, the code which implements the proposed methodology has been added to the existing R software package "$ICsurv$" and is freely available from the CRAN.

Briefly, the methodology offers modeling flexibility through the use of monotone splines [64] for estimating the cumulative baseline hazard. To estimate all the unknown parameters, a computationally efficient EM algorithm is developed through a data augmentation scheme involving latent Poisson random variables. At each iteration of this algorithm, the spline coefficients are updated in closed form, with the regression parameters being updated through solving a low-dimensional system of equations. Through extensive simulation studies [89], the proposed methodology is shown to provide reliable estimation and inference with respect to the covariate effects, baseline survival function, and baseline probability of experiencing an instantaneous failure. The in-depth discussion of this approach is provided in Chapter 3. In Chapter 4, we assess performance and power of this methodology in context of the data originating from the clinical trial.

2.3.2 SU Degradation Modeling Based on a Time-Dependent Ornstein Uhlenbeck Process

In this analysis, we extend the scope of efficacy endpoint to include the time to loss of SU to any arbitrary threshold, for e.g., expected population threshold reducing by half (also called "$half - life$" of population sensitivity threshold) or individual's expected threshold dropping by at least a level of the administered dose or even below MCRT. To this end, we consider modeling individual threshold trajectories to emulate a systematic biological immune process that is expected upon therapy cessation. This entails modeling a process-governing systematic trajectory that constitutes repeated threshold measurements over time with perturbations due to individual differences; it is reasonable to assume

that responses for all subjects follow the same biological process, but with varying initial thresholds and random fluctuations among individuals.

Dynamical biological processes are often described by systems of stochastic differential equations that account for noisy components often present in biological systems. The system noise responsible for internal system uncertainties is modeled by including a diffusion term in the driving equations. On other hand, external stochastic impacts such as inter-individual variability is modeled using mixed-effects framework that allows splitting total variation into its within- and between-individual components, leading to more precise estimation of population parameters. This inter-individual variability is modeled with the random effect, and the intra-individual variability with an additive noise term. When both system noise and random effects are considered, stochastic differential mixed-effects models emerge [19]. This class of models, where random effects are incorporated into a SDE model, enable simultaneous depiction of system noise in the process dynamics and variability between experimental units, thus providing a powerful and flexible modeling tool suited to biomedical applications involving pharmaco-kinetic/pharmaco-dynamic studies [63; 79].

SDEs are adopted widely in the field of reliability engineering to model degradation processes with temporary fluctuations in an overall degrading trend [18]. But because of the complex parameter estimation in SDE models, only a few cases can be solved explicitly using likelihood inference with explicit transition density. The Ornstein-Uhlenbeck (OU) [45; 81] process is one of few cases which can be solved explicitly with also explicit probability laws in SDEs. However, application of such a process into the field of pharmacodynamics could be an interesting but challenging balance between the modeling and the computability. Our concentration in this thesis will be put on extending and estimating the time-dependent OU process given by $dY_t = [A(t)Y_t + B(t)]dt + \sigma(t)dW_t$, where dY_t represents change in SU threshold over a finite time interval, $A(t)$ is the process drift, $B(t)$ explains reverting mean, $\sigma(t)$ represents process diffusion.

Taking into account expert clinical opinion, we justify the use of OU based diffusion process for modeling SU degradation trajectories with the following assumptions: (i) in the absence of therapy, individual's threshold trajectory that on an average is monotonically non-increasing over time, will have fluctuations over relatively short time intervals; (ii) thresholds are log normally distributed to have normally distributed threshold change. This allows to measure percentage change in threshold rather than absolute change, since the change is likely to depend on the initial value; (iii) individual

will revert to the baseline threshold in the absence of immunotherapy, or perhaps have a slight boost in the immune system forever; (iv) population mean will decay exponentially at a rate different from the drift rate of an individual; (v) finally, individuals differ in their inherent tendency to lose their SU threshold over time. Thus, the model assumes that the process dynamics responsible for evolving threshold in each subject follows a common functional form, and the inter-individual differences arise due to different realizations of the Brownian motion paths and of random parameters.

A variety of statistical inference methods for discretely observed system diffusion processes have been developed during the past decades [20; 3]. On other hand, the theory for more commonly applied mixed-effects models, both linear and non-linear, is also well developed for deterministic models (without system error) [9; 84], and standard software for model fitting is available [15; 49]. However, to our knowledge there is limited theory at present for SDE models with random effects. [59] implemented SDEs in a non-linear mixed-e ects model, where parameter estimation was performed by an approximation of the likelihood function that is constructed by combining the first-order conditional estimation (FOCE) method with the extended Kalman filter. Recently, [22] considered non-linear mixed models defined by SDEs wherein the parameters of the diffusion process have random effects and proposed a maximum likelihood estimation method based on the stochastic approximation EM algorithm. As SDE models are gaining more popular venue in biomedical applications, there is a pressing need for developing mixed effects theory and parameter estimation. [39] reviews methods for PK/PD modeling, but regrets that system noise is not considered due to difficult estimation and lack of software in the pharmacokinetic field.

In this work, we present OU based SDMEM and derive an expression for the likelihood function and transition density in context of the interval-censored data emanating from the study. Inspired by the the works of [88; 90], we adopt Bayesian framework to estimate the process parameters using hybrid Metropolis-Hastings algorithm within a Gibbs sampler and assess finite sample performance of this procedure based on extensive simulation studies. Once the process parameters are estimated, we propose to apply direct maximization techniques or use Bayesian framework to estimate FPT to threshold levels of interest defined at the beginning of this section.

CHAPTER 3: A PROPORTIONAL HAZARDS MODEL FOR INTERVAL-CENSORED DATA SUBJECT TO INSTANTANEOUS FAILURES

The proportional hazards (PH) model is one of the most popular models used to analyze time to event data arising from clinical trials and longitudinal studies, among others. In many such studies, the event time of interest is not directly observed but rather is known relative to periodic examination times; i.e., practitioners observe either current status or interval-censored data. The analysis of data of this structure is often fraught with many difficulties. Further exacerbating this issue, in some such studies the observed data also consists of instantaneous failures; i.e., the event times for several study units coincide exactly with the time at which the study begins. In light of these difficulties, this work focuses on developing a mixture model, under the PH assumptions, which can be used to analyze interval-censored data subject to instantaneous failures. To allow for modeling flexibility, two methods of estimating the unknown cumulative baseline hazard function are proposed; a fully parametric and a monotone spline representation are considered. Through a data augmentation procedure involving latent Poisson random variables, an expectation-maximization (EM) algorithm is developed to complete model fitting. The resulting EM algorithm is easy to implement and is computationally efficient. Moreover, through extensive simulation studies the proposed approach is shown to provide both reliable estimation and inference. The motivation for this work arises from an ongoing randomized clinical trial funded by the National Institutes of Allergy and Infectious Diseases, aimed at assessing the effectiveness of a new peanut allergen treatment with respect to desensitization in children.

3.1 Introduction

Interval-censored data commonly arise in many clinical trials and longitudinal studies, and is characterized by the fact that the event time of interest is not directly observable, but rather is known relative to observation times. As a special case, current status data (or case-1 interval censoring)

arise when there exists exactly one observation time per study unit; i.e., at the observation time one discovers whether or not the event of interest has occurred. Data of this structure often occurs in resource limited environments or due to destructive testing. Alternatively, general interval-censored data (or case-2 interval censoring) arise when multiple observation times are available for each study unit, and the event time can be ascertained relative to two observation times. It is well known that ignoring the structure of interval-censored data during an analysis can lead to biased estimation and inaccurate inference; see [58; 23]. Further exasperating this issue, some studies are subject to the occurrence of instantaneous failures; i.e., the event time of interest for a number of the study units occurs at time zero. This feature can occur as an artifact of the study design or may arise during an intent-to-treat analysis [50; 52; 44]. For example, [16] describes a registry based study of end-stage renal disease patients, with the time of enrollment corresponding to the time at which the patient first received dialysis. In this study, several of the patients expire during the first dialysis treatment, leading to the occurrence of an instantaneous failure. Survival data with instantaneous events is not uncommon in epidemiological and clinical studies, and for this reason, herein a general methodology under the proportional hazards (PH) model is developed for the analysis of interval-censored data subject to instantaneous failures.

Originally proposed by [17], the PH model has (arguably) become one of the most popular regression models for analyzing time-to-event data. For analyzing interval-censored data under the PH model, several notable contributions have been made in the recent years; e.g., see [25; 69; 28; 32; 60; 61; 27; 7; 14; 92; 38; 94; 47]. More recently, [86] developed a methodology under the PH model which can be used to accurately and reliably analyze interval-censored data. In particular, this approach makes use of a monotone spline representation to approximate the cumulative baseline hazard function. In doing so, an expectation-maximization (EM) algorithm is developed through a data augmentation scheme involving latent Poisson random variables which can be used to complete model fitting. It is worthwhile to note, that none of the aforementioned techniques were designed to account for the effects associated with instantaneous failures.

The phenomenon of instantaneous (or early) failures occur in many lifetime experiments; to include, but not limited to, reliability studies and clinical trials. In reliability studies, instantaneous failures may be attributable to inferior quality or faulty manufacturing, where as in clinical trials these events may manifest due to adverse reactions to treatments or clinical definitions of outcomes.

When the failure times are exactly observed, as is the case in reliability studies, it is common to incorporate instantaneous failures through a mixture of parametric models, with one being degenerate at time zero; e.g., see [54; 40; 57; 55; 62; 42]. In the case of interval-censored data, more common among epidemiological studies and clinical trials, accounting for instantaneous failures becomes a more tenuous task, with practically no guidance available among the existing literature. Arguably, in the context of interval-censored data, one could account for instantaneous failures by introducing an arbitrarily small constant for each as an observation time, and subsequently treat the instantaneous failures as left-censored observations. In doing so, methods for interval-censored data, such as those discussed above, could be employed. While this approach may seem enticing, in the case of a relatively large number of instantaneous failures it has several pitfalls. In particular, through numerical studies (results not shown) it has been determined that this approach when used in conjunction with modeling techniques such as those proposed in [60; 86] may lead to inaccurate estimation of the survival curves and/or the covariate effects. Further, after an extensive literature review, it does not appear that any methodology has previously been developed to specifically address data of this structure. For these reasons, herein a general methodology under the PH model is developed for the analysis of interval-censored data subject to instantaneous failures.

This work seeks to provide a comprehensive methodology that can be utilized to analyze interval-censored data subject to instantaneous failures. In particular, several primary contributions are made: (1) a new mixture model is proposed, (2) a computationally efficient algorithm is developed for fitting the proposed model, and (3) a technique for uncertainty quantification is outlined. The new mixture model, which is a generalization of the semi-parametric PH model studied in [86], is developed under the standard PH assumption; i.e., the covariates provide for a multiplicative effect on the baseline risk of experiencing a failure both at time zero and thereafter. Two separate techniques are developed for the purposes of estimating the cumulative baseline hazard function. The first allows a practitioner to specify a parametric form (up to a collection of unknown coefficients) for the unknown function, while the second provides for more modeling flexibility through the use of the monotone splines of [64]. Under either formulation, a two-stage data augmentation scheme involving latent Poisson random variables is used to develop an efficient EM algorithm which can be used to estimate all of the unknown parameters. Through extensive simulation studies the proposed methodology is shown to provide reliable estimation and inference with respect to

the covariate effects, baseline cumulative hazard function, and baseline probability of experiencing an instantaneous failure. This work is primarily motivated by an ongoing randomized clinical trial supported by the National Institutes of Allergy and Infectious Diseases (NIAID) being conducted at the University of North Carolina at Chapel Hill, and is aimed at developing, assessing, and validating the proposed approach as a viable tool which can be used to analyze the data resulting from this trial once it is complete.

The remainder of this article is organized as follows. In Section 2, an in depth discussion of the motivating randomized clinical trial is provided. Section 3 presents the development of the proposed model, the derivation of the EM algorithm, and outlines uncertainty quantification. The finite sample performance of the proposed approach is evaluated through extensive numerical studies, the features and results of which are provided in Section 4. Section 5 concludes with a summary discussion. Further, code which implements the proposed methodology has been added to the existing R software package `ICsurv` and is freely available from the CRAN (i.e., http://cran.us.rproject.org/).

3.2 Motivating Application

Supported by the National Institutes of Allergy and Infectious Diseases and conducted at the University of North Carolina at Chapel Hill, the Peanut Sublingual Immunotherapy (SLIT) and Induction of Clinical Tolerance in Peanut Allergic Children is an ongoing clinical trial that was initiated in 2012 to assess the effectiveness of peanut SLIT to induce clinical tolerance. The protocol was revised in 2016 with an altered study design to assess time to loss of tolerance among subjects desensitized to peanut allergen after 48 months of peanut SLIT induction. Participants include children between 1-11 years of age at the time of enrollment. The revised study, still underway, consists of a build-up/maintenance phase (approximately, 48 months), wherein SLIT therapy is incremented during the initial 6 months and maintained thereafter, and a tolerance phase (17 weeks), wherein the therapy is discontinued. Each phase ends with a Double-Blind Placebo Control Food Challenge (DBPCFC), a gold standard, to assess desensitization- (an increase in reaction threshold while receiving peanut SLIT) and loss of tolerance- (maintaining an increase in reaction threshold after discontinuing study drug) thresholds. The tolerance phase consists of staggered examination time-points between 1-17 weeks that each subject has an equal probability of being randomized to for

administration of final DBPCFC. Further details on the study design can be found in the companion paper [M et al.].

The study underway has a well defined end point, the time to loss of sustained unresponsiveness (SU) to a targeted dose level, that is measured as the time from the 48th month DBPCFC (time of study entry) until the final DBPCFC. Due to the staggered randomized time-points of the final DBPCFC, the endpoints are not observable but rather are known relative to the times of the final DBPCFC, resulting in case-1 interval-censored data, with non-informative censoring times. Further, a few subjects that either fail to meet a minimal clinically relevant threshold (MCRT) or exhibit a threshold lower than the targeted dose level at the study entry, result in instantaneous events with some probability mass at time zero. At enrollment, several risk factors were collected on each patient and include age, gender, and several clinical measurements. The primary objectives of the analysis of the data arising from this study, once completed, involves assessing the association of risk factors with the end point, as well as estimating the baseline survival function and the instantaneous failure probability.

3.3 Model and Methodology

Let T denote the failure time of interest. Under the PH model, the survival function can be generally written as

$$S(t|\mathbf{x}) = S_0(t)^{e^{\mathbf{x}'\boldsymbol{\beta}}} \tag{3.1}$$

where $\mathbf{x}$ is a $(r \times 1)$-dimensional vector of covariates, $\boldsymbol{\beta}$ is the corresponding vector of regression coefficients, and $S_0(t)$ is the baseline survival function. Under the phenomenon of interest, there is a baseline risk (probability) of experiencing an instantaneous failure; i.e., $S(0|\mathbf{x} = \mathbf{0}_r) = S_0(0) = 1-p$, where $p \in [0, 1]$ is the baseline risk and $\mathbf{0}_r$ is a $(r \times 1)$-dimensional vector of zeros. Thus, under the PH assumptions, the probability of experiencing an instantaneous failure, given the covariate information contained in $\mathbf{x}$, can be ascertained from (3.1) as

$$\begin{aligned} P(T = 0|\mathbf{x}) &= 1 - S(0|\mathbf{x}) \\ &= 1 - (1-p)^{e^{\mathbf{x}'\boldsymbol{\beta}}}. \end{aligned}$$

Similarly, given that an instantaneous failure does not occur, it is assumed that the failure time conditionally follows the standard PH model; i.e.,

$$P(T > t|\mathbf{x}, T > 0) = 1 - F(t|\mathbf{x}),$$

where $F(t|\mathbf{x}) = 1 - \exp\{-\Lambda_0(t)\exp(\mathbf{x}'\boldsymbol{\beta})\}$ and $\Lambda_0(\cdot)$ is the usual cumulative baseline hazard function. Note, in order for $F(\cdot|\mathbf{x})$ to be a proper cumulative distribution function, $\Lambda_0(\cdot)$ should be a monotone increasing function with $\Lambda_0(0) = 0$. Thus, through an application of the Law of Total Probability, one has that

$$\begin{aligned} P(T > t|\mathbf{x}) &= P(T > t|\mathbf{x}, T > 0)P(T > 0|\mathbf{x}) \\ &= \{1 - F(t|\mathbf{x})\}(1-p)^{e^{\mathbf{x}'\boldsymbol{\beta}}}, \end{aligned}$$

for $t > 0$. Based on these assumptions, the cumulative distribution function of T can be expressed as the following mixture model,

$$H(t|\mathbf{x}) = \begin{cases} 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}, & \text{for } t = 0, \\ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}\{1 - F(t|\mathbf{x})\}, & \text{for } t > 0, \end{cases}$$

where, for reasons that will shortly become apparent, $1 - p$ is re-parametrized as $\exp(-\alpha)$, for $\alpha > 0$.

3.3.1 Observed data likelihood

In scenarios where interval-censored data arise, one has that the failure time (T) is not directly observed, but rather is known relative to two observation times, say $L < R$; i.e., one has that $L < T < R$. In general, the four different outcomes considered here can be represented through the values of L and R; i.e., an instantaneous failure ($L = R = 0$), T is left-censored ($0 = L < R < \infty$), T is interval-censored ($0 < L < R < \infty$), and T is right-censored ($0 < L < R = \infty$). For notational convenience, let ψ be an indicator denoting the event that T is not an instantaneous failure, and δ_1, δ_2, and δ_3 be censoring indicators denoting left-, interval-, and right-censoring,

respectively; i.e., $\psi = I(T > 0)$, $\delta_1 = I(0 = L < R < \infty)$, $\delta_2 = I(0 < L < R < \infty)$, and $\delta_3 = (0 < L < R = \infty)$.

In order to derive the observed data likelihood, it is assumed throughout that the individuals are independent, and that conditional on the covariates, the failure time for an individual is independent of the observational process. This assumption is common among the survival literature; see, e.g., [94; 51] and the references therein. The observed data collected on n individuals is given by $\boldsymbol{D} = \{(L_i, R_i, \mathbf{x}_i, \psi_i, \delta_{i1}, \delta_{i2}, \delta_{i3}); i = 1, 2, \ldots, n\}$, which constitutes n independent realization of $\{(L, R, \mathbf{x}, \psi, \delta_1, \delta_2, \delta_3)$. Thus, under the aforementioned assumptions, the observed data likelihood is given by

$$\begin{aligned} L_{obs}(\boldsymbol{\theta}) &= \prod_{i=1}^{n} \Big[F(R_i|\mathbf{x}_i)^{\delta_{i1}} \{F(R_i|\mathbf{x}_i) - F(L_i|\mathbf{x}_i)\}^{\delta_{i2}} \{1 - F(L_i|\mathbf{x}_i)\}^{\delta_{i3}} \Big]^{\psi_i} \\ &\quad \Big\{ e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} \Big\}^{\psi_i} \Big\{ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} \Big\}^{1-\psi_i}, \end{aligned} \tag{3.2}$$

where $\boldsymbol{\theta}$ represents the set of unknown parameters which are to be estimated.

3.3.2 Representations of $\Lambda_0(\cdot)$

The unknown parameters in the observed likelihood involve the regression parameters $\boldsymbol{\beta}$, α, and the cumulative baseline hazard function $\Lambda_0(\cdot)$. Herein, two techniques for modeling the cumulative baseline hazard function are discussed. The first approach considers the use of a fully parametric model, which is known up to a set of unknown coefficients. For example, a linear, quadratic, or logarithmic parametric model can be specified by setting $\Lambda_0(t) = \gamma_1 t$, $\Lambda_0(t) = \gamma_1 t + \gamma_2 t^2$, and $\Lambda_0(t) = \gamma_1 \log(1 + t)$, respectively. Note, all of these models obey the constraints placed on $\Lambda_0(\cdot)$, as long as the $\gamma_l > 0$, for $l = 1, 2$. In general, a parametric form for the cumulative baseline hazard model can be specified as

$$\Lambda_0(t) = \sum_{l=1}^{k} \gamma_l b_l(t), \tag{3.3}$$

where $b_l(\cdot)$ is a monotone increasing function, $b_l(0) = 0$, and $\gamma_l > 0$, for $l = 1, ..., k$. Under these mild conditions, it is easily verified that $\Lambda_0(\cdot)$ inherits the same properties, and therefore adheres to the aforementioned constraints.

The second approach, which is inspired by the works of [86; 48; 85; 13; 53; 87], views $\Lambda_0(\cdot)$ as an unknown function, and hence an infinite dimensional parameter. To reduce the dimensionality of the problem, the monotone splines of [64] are used to approximate $\Lambda_0(\cdot)$. Structurally, this representation is identical to that of (3.3) with the exception that $b_l(\cdot)$ is a spline basis function and γ_l is an unknown spline coefficient, for $l = 1, \ldots, k$. Again, it is required that $\gamma_l > 0$, for all l, to ensure that $\Lambda_0(\cdot)$ is monotone increasing function. Briefly, the spline basis functions are piecewise polynomial functions and are fully determined once a knot sequence and the degree are specified. The shape of the basis splines are predominantly determined by the placement of the knots while the degree controls the smoothness [13]. For instance, specifying the degree to take values 1, 2 or 3 correspond to the use of linear, quadratic or cubic polynomials, respectively. Given the m knots and degree, the k ($k = m+$ degree $- 2$) basis functions are fully determined. For further discussion on specifying the knots, as well as their placement, see [86; 64; 53].

3.3.3 Data Augmentation

Under either of the representations of $\Lambda_0(\cdot)$ proposed in Section 3.2, the unknown parameters in the observed data likelihood consist of $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \alpha)'$, where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_k)'$. Since the observed data likelihood exists in closed-form, the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ could be obtained by directly maximizing (3.2) with respect to $\boldsymbol{\theta}$; i.e., one could obtain $\hat{\boldsymbol{\theta}}$, the MLE of $\boldsymbol{\theta}$, as $\hat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} L_{obs}(\boldsymbol{\theta})$. It is worthwhile to point out that the numerical process of directly maximizing (3.2), with respect to $\boldsymbol{\theta}$, is often unstable and rarely performs well [87].

To circumvent these numerical instabilities, an EM algorithm was derived for the purposes of identifying the MLE. This algorithm was developed based on a two-stage data augmentation process, where carefully structured latent Poisson random variables are introduced as missing data. The first stage relates both the instantaneous failure indicator and the censoring indicators to latent Poisson random variables; i.e., the Z_i, W_i, and Y_i are introduced such that

$$\begin{aligned} Z_i &\sim \text{Poisson}\{\Lambda_0(t_{i1}) \exp(\mathbf{x}_i'\boldsymbol{\beta})\}, \\ W_i &\sim \text{Poisson}[\{\Lambda_0(t_{i2}) - \Lambda_0(t_{i1})\} \exp(\mathbf{x}_i'\boldsymbol{\beta})], \\ Y_i &\sim \text{Poisson}\{\alpha \exp(\mathbf{x}_i'\boldsymbol{\beta})\}, \end{aligned}$$

subject to the following constraints: $\delta_{i1} = I(Z_i > 0)$, $\delta_{i2} = I(Z_i = 0, W_i > 0)$, $\delta_{i3} = I(Z_i = 0, W_i = 0)$, and $\psi_i = I(Y_i = 0)$, where $t_{i1} = R_i I(\delta_{i1} = 1) + L_i I(\delta_{i1} = 0)$, and $t_{i2} = R_i I(\delta_{i2} = 1) + L_i I(\delta_{i3} = 1)$. At this stage of the data augmentation, the conditional likelihood is

$$L_A(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left\{ P_{Z_i}(Z_i) P_{W_i}(W_i)^{\delta_{i2}+\delta_{i3}} C_i \right\}^{\psi_i} P_{Y_i}(Y_i) I(Y_i = 0)^{\psi_i} I(Y_i > 0)^{(1-\psi_i)}, \tag{3.4}$$

where $C_i = \delta_{i1} I(Z_i > 0) + \delta_{i2} I(Z_i = 0, W_i > 0) + \delta_{i3} I(Z_i = 0, W_i = 0)$ and $P_A(\cdot)$ is the probability mass function of the random variable A. In the second and final stage, the Z_i and W_i are separately decomposed into the sum of k independent latent Poisson random variables; i.e., $Z_i = \sum_{l=1}^{k} Z_{il}$ and $W_i = \sum_{l=1}^{k} W_{il}$, where

$$\begin{aligned} Z_{il} &\sim \text{Poisson}\{\gamma_l b_l(t_{i1}) \exp(\mathbf{x}_i' \boldsymbol{\beta})\}, \\ W_{il} &\sim \text{Poisson}[\{\gamma_l b_l(t_{i2}) - \gamma_l b_l(t_{i1})\} \exp(\mathbf{x}_i' \boldsymbol{\beta})]. \end{aligned}$$

At this stage, the augmented data likelihood is

$$\begin{aligned} L_C(\boldsymbol{\theta}) \;=\; & \prod_{i=1}^{n} \prod_{l=1}^{k} \Big[P_{Z_{il}}(Z_{il}) I(Z_i = Z_{i\cdot}) \{ P_{W_{il}}(W_{il}) I(W_i = W_{i\cdot}) \}^{\delta_{i2}+\delta_{i3}} C_i \Big]^{\psi_i} \\ & P_{Y_i}(Y_i) I(Y_i = 0)^{\psi_i} I(Y_i > 0)^{(1-\psi_i)}, \end{aligned} \tag{3.5}$$

where $Z_{i\cdot} = \sum_{l=1}^{k} Z_{il}$ and $W_{i\cdot} = \sum_{l=1}^{k} W_{il}$. It is relatively easy to show that by integrating (3.5) over the latent random variables one will obtain the observed data likelihood depicted in (3.2).

3.3.4 EM algorithm

In general, the EM algorithm consists of two steps: the expectation step (E-step) and the maximization step (M-step). The E-step in this algorithm involves taking the expectation of $\log\{L_c(\boldsymbol{\theta})\}$ with respect to all latent variables conditional on the current parameter value $\boldsymbol{\theta}^{(d)} = (\boldsymbol{\beta}^{(d)\prime}, \boldsymbol{\gamma}^{(d)\prime}, \alpha^{(d)})'$ and the observed data $\boldsymbol{D}$. This results in obtaining the $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ function, where $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)}) = E[\log\{L_c(\boldsymbol{\theta})\} | \boldsymbol{D}, \boldsymbol{\theta}^{(d)}]$. The M-step then finds $\boldsymbol{\theta}^{(d+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$. This process is repeated in turn until convergence of the algorithm is attained. In this particular setting,

the E-step yields $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ as

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)}) &= \sum_{i=1}^{n}\sum_{l=1}^{k} \psi_i \Big[\{E(Z_{il}) + (\delta_{i2} + \delta_{i3})E(W_{il})\}\{\log(\gamma_l) + \mathbf{x}_i'\boldsymbol{\beta}\} \\ &\qquad -\gamma_l e^{\mathbf{x}_i'\boldsymbol{\beta}}\{(\delta_{i2} + \delta_{i1})b_l(R_i) + \delta_{i3}b_l(L_i)\}\Big] \\ &\quad + \sum_{i=1}^{n} E(Y_i)\log(\alpha e^{\mathbf{x}_i'\boldsymbol{\beta}}) - \alpha e^{\mathbf{x}_i'\boldsymbol{\beta}} + H(\boldsymbol{\theta}^{(d)}), \end{aligned}$$

where $H(\boldsymbol{\theta}^{(d)})$ is a function of $\boldsymbol{\theta}^{(d)}$ but is free of $\boldsymbol{\theta}$. Notice that in $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ we suppress, for notational convenience, the dependence of the expectations on the observed data and $\boldsymbol{\theta}^{(d)}$; i.e., from henceforth it should be understood that $E(\cdot) = E(\cdot|\boldsymbol{D}, \boldsymbol{\theta}^{(d)})$.

An enticing feature, which makes the proposed approach computationally efficient, is that all of the expectations in $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ can be expressed in closed-form, and moreover can be computed via simple matrix and vector operations. In particular, from (3.4) it can be ascertained that if $\delta_{i1} = 1$ and $\psi_i = 1$ then Z_i conditionally, given $\boldsymbol{\theta}^{(d)}$ and $\boldsymbol{D}$, follows a zero-truncated Poisson distribution, and it follows a degenerate distribution at 0 for any other values of δ_{i1} and ψ_i. Thus, the conditional expectation of Z_i, given $\boldsymbol{\theta}^{(d)}$ and $\boldsymbol{D}$, can be expressed as

$$E(Z_i) = \delta_{i1}\psi_i\Lambda_0^{(d)}(t_{i1})\exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})\left[1 - \exp\{-\Lambda_0^{(d)}(t_{i1})\exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})\}\right]^{-1},$$

where $\Lambda_0^{(d)}(t) = \sum_{l=1}^{k}\gamma_l^{(d)}b_l(t)$. Through a similar set of arguments one can obtain the necessary conditional expectations of W_i and Y_i as

$$\begin{aligned} E(W_i) &= \delta_{i2}\psi_i\{\Lambda_0^{(d)}(t_{i2}) - \Lambda_0^{(d)}(t_{i1})\}\exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)}) \\ &\qquad \Big(1 - \exp[-\{\Lambda_0^{(d)}(t_{i2}) - \Lambda_0^{(d)}(t_{i1})\}\exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})]\Big)^{-1}, \\ E(Y_i) &= (1-\psi_i)\alpha^{(d)}\exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})\left[1 - \exp\{-\alpha^{(d)}\exp(\mathbf{x}_i'\boldsymbol{\beta}^{(d)})\}\right]^{-1}, \end{aligned}$$

respectively. Further, from (3.5) it can be ascertained that if $\delta_{i1} = 1$ and $\psi_i = 1$ then Z_{il} conditionally, given $Z_i, \boldsymbol{D}$ and $\boldsymbol{\theta}^{(d)}$, follows a binomial distribution with Z_i being the number of trials and $\gamma_l^{(d)}b_l(t_{i1})\{\Lambda_0^{(d)}(t_{i1})\}^{-1}$ being the success probability, and it is follows a degenerate distribution at 0 for any other values of δ_{i1} and ψ_i. Thus, through an application of the Law of Iterated Expectations,

the conditional expectation of Z_{il}, given $\boldsymbol{\theta}^{(d)}$ and $\boldsymbol{D}$, can be expressed as

$$E(Z_{il}) = E(Z_i)\gamma_l^{(d)} b_l(t_{i1})\{\Lambda_0^{(d)}(t_{i1})\}^{-1}.$$

Through a similar set of arguments one can obtain the necessary conditional expectation of W_{il} as

$$E(W_{il}) = E(W_i)\gamma_l^{(d)}\{b_l(t_{i2}) - b_l(t_{i1})\}\{\Lambda_0^{(d)}(t_{i2}) - \Lambda_0^{(d)}(t_{i1})\}^{-1}.$$

Note, in the expressions of the expectations of Z_{il} and W_{il} the dependence on δ_{i1}, δ_{i2}, and ψ_i are suppressed with the properties associated with these variables being inherited from the expectations associated with Z_i and W_i, respectively.

The M-step of the algorithm then finds $\boldsymbol{\theta}^{(d+1)} = \text{argmax}_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$. To this end, consider the partial derivatives of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ with respect to $\boldsymbol{\theta}$ which are given by

$$\begin{aligned}
\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})}{\partial \gamma_l} &= \sum_{i=1}^{n} \psi_i \big[\gamma_l^{-1}\{E(Z_{il}) + (\delta_{i2} + \delta_{i3})E(W_{il})\} - e^{\mathbf{x}_i'\boldsymbol{\beta}}\{(\delta_{i2} + \delta_{i1})b_l(R_i) \\
&\qquad + \delta_{i3} b_l(L_i)\}\big], \qquad (3.6) \\
\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})}{\partial \alpha} &= \sum_{i=1}^{n} -e^{\mathbf{x}_i'\boldsymbol{\beta}} + \alpha^{-1}E(Y_i), \qquad (3.7) \\
\frac{\partial Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^{n} \big[\psi_i\{E(Z_i) + \delta_{i2}E(W_i)\} - \psi_i\{(\delta_{i1} + \delta_{i2})\Lambda_0(R_i) + \delta_{i3}\Lambda_0(L_i)\}e^{\mathbf{x}_i'\boldsymbol{\beta}} \\
&\qquad - \alpha e^{\mathbf{x}_i'\boldsymbol{\beta}} + E(Y_i)\big]\mathbf{x}_i. \qquad (3.8)
\end{aligned}$$

By setting (3.6) equal to zero and solving for γ_l, one can obtain

$$\gamma_l^*(\boldsymbol{\beta}) = \frac{\sum_{i=1}^{n} \psi_i\{E(Z_{il}) + \delta_{i2}E(W_{il})\}}{\sum_{i=1}^{n} \psi_i\{(\delta_{i2} + \delta_{i1})b_l(R_i) + \delta_{i3}b_l(L_i)\}e^{\mathbf{x}_i'\boldsymbol{\beta}}}, \qquad (3.9)$$

for $l = 1, \ldots, k$. Similarly, by setting (3.7) equal to zero and solving for α, one can obtain

$$\alpha^*(\boldsymbol{\beta}) = \frac{\sum_{i=1}^{n} E(Y_i)}{\sum_{i=1}^{n} e^{\mathbf{x}_i'\boldsymbol{\beta}}}. \qquad (3.10)$$

Notice that $\gamma_l^*(\boldsymbol{\beta})$ and $\alpha^*(\boldsymbol{\beta})$ depend on $\boldsymbol{\beta}$. Thus, one can obtain $\boldsymbol{\beta}^{(d+1)}$ by setting (3.8) equal to zero and solving the resulting system of equations for $\boldsymbol{\beta}$, after replacing γ_l and α by $\gamma_l^*(\boldsymbol{\beta})$ and $\alpha^*(\boldsymbol{\beta})$, respectively. Note, the aforementioned system of equations can easily be solved using a standard Newton Raphson approach. Finally, one obtains $\gamma_l^{(d+1)}$ and $\alpha^{(d+1)}$ as $\gamma_l^*(\boldsymbol{\beta}^{(d+1)})$ and $\alpha^*(\boldsymbol{\beta}^{(d+1)})$, respectively.

The proposed EM algorithm is now succinctly stated. First, initialize $\boldsymbol{\theta}^{(0)}$ and repeat the following steps until converges.

1. Calculate $\boldsymbol{\beta}^{(d+1)}$ by solving the following system of equations

$$\sum_{i=1}^{n} \left[\psi_i\{E(Z_i) + \delta_{i2}E(W_i)\} - \alpha^*(\boldsymbol{\beta})e^{\mathbf{x}_i'\boldsymbol{\beta}} + E(Y_i)\right]\mathbf{x}_i = \sum_{i=1}^{n}\sum_{l=1}^{k} \psi_i\{(\delta_{i1} + \delta_{i2})b_l(R_i) + \delta_{i3}b_l(L_i)\}\gamma_l^*(\boldsymbol{\beta})e^{\mathbf{x}_i'\boldsymbol{\beta}}\mathbf{x}_i,$$

   where $\gamma_l^*(\boldsymbol{\beta})$ and $\alpha^*(\boldsymbol{\beta})$ are defined above.

2. Calculate $\gamma_l^{(d+1)} = \gamma_l^*\left(\boldsymbol{\beta}^{(d+1)}\right)$ for $l = 1, \ldots, k$ and $\alpha^{(d+1)} = \alpha^*\left(\boldsymbol{\beta}^{(d+1)}\right)$.

3. Update $d = d + 1$.

At the point of convergence, define $\boldsymbol{\theta}^{(d)} = (\boldsymbol{\beta}^{(d)'}, \boldsymbol{\gamma}^{(d)'}, \alpha^{(d)})'$ to be the proposed estimator $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\gamma}}', \hat{\alpha})'$, which is the MLE of $\boldsymbol{\theta}$.

3.3.5 Variance estimation

For the purposes of uncertainty quantification, several variance estimators were considered and evaluated; e.g., the inverse of the observed Fisher information matrix, the Huber sandwich estimator, and the outer product of gradients (OPG) estimator. After extensive numerical studies (results not shown), it was found that the most reliable variance estimator, among those considered, was that of the OPG estimator. In general, the OPG estimator is obtained as

$$\widehat{V}(\widehat{\boldsymbol{\theta}}) = \left[\frac{1}{n}\sum_{i=1}^{n} \hat{g}_i(\widehat{\boldsymbol{\theta}})\hat{g}_i'(\widehat{\boldsymbol{\theta}})\right]^{-1}$$

where $\hat{g}_i(\widehat{\boldsymbol{\theta}}) = \partial l_i(\boldsymbol{\theta})/\partial\boldsymbol{\theta}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$ and $l_i(\boldsymbol{\theta})$ is the log-likelihood contribution of the ith individual. Using this estimator, one can conduct standard Wald type inference.

3.4 Simulation Study

In order to investigate the finite sample performance of the proposed methodology, the following simulation study was conducted. The true distribution of the failure times was specified to be

$$H(t|\mathbf{x}) = \begin{cases} 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}, & \text{for } t = 0, \\ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}\{1 - F(t|\mathbf{x})\}, & \text{for } t > 0, \end{cases}$$

where $p = 0.3$ (i.e., $\alpha = -\log(0.7)$), $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2)'$, $\mathbf{x}_1 \sim \text{N}(0, 1)$, $\mathbf{x}_2 \sim \text{Bernoulli}(0.5)$, and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, where β_1 and β_2 take on values of -0.5 and 0.5 resulting 4 regression parameter configurations. Additionally, these studies consider two cumulative baseline hazard functions; i.e., a logarithmic $\Lambda_0(t) = \log(t + 1)/\log(11)$ and a linear $\Lambda_0(t) = 0.1t$. These choices were made so that the hazard functions behave similarly but have different shapes. In total, these specifications lead to eight separate data generating models for the failure times. Two generating processes were considered for the observation times: an exponential distribution with a mean of 10 and a discrete uniform over the interval $[1, 17]$. In both cases, a single observation time, O, was generated for each failure time which was not instantaneous (i.e., $T > 0$), and the intervals were created such that $L = 0$ ($R = \infty$) and $R = O$ ($L = O$) if T was smaller (greater) than O. A few comments are warranted on the selection of the observation processes. First, the latter process is actually indicative of the observation process considered in the motivating clinical trial, while the former attempts to match the baseline characteristics of the failure time distribution. Second, note that the specification of the two observation processes result in case-1 interval-censored (i.e., current status) data. This was done for two primary reasons: first, data of this structure is going to be collected as a part of the motivating clinical trial, and second, data of this nature possess less information when compared to general interval-censored data, and thus if the proposed approach works well in this setting it could be inferred that it should perform better in the case of interval-censored data. In total, these data generating steps lead to sixteen generating mechanisms, and each were used to create 500 independent data sets, each consisting of $n = 100$ observations.

In order to examine the performance of the proposed approach across a broad spectrum of characteristics, several different models were used to analyze each data set. First, following the development presented in Section 3.2, three different parametric forms were considered for the cumulative baseline hazard function: $\Lambda_{0_1}(t) = \gamma_1 \log(t+1)$, $\Lambda_{0_2}(t) = \gamma_1 t$, and $\Lambda_{0_3}(t) = \gamma_1 t + \gamma_2 t^2$, which are henceforth referred to as models M1, M2, and M3, respectively. Note, these specifications allow one to examine the performance of the proposed approach when the cumulative baseline hazard function is correctly specified (e.g., M2 when $\Lambda_0(t) = 0.1t$), over specified (e.g., M3 when $\Lambda_0(t) = 0.1t$), and misspecified (e.g., M1 when $\Lambda_0(t) = 0.1t$). Further, for each data set a model (M4) was fit using the monotone spline representation for the cumulative baseline hazard function developed in Section 3.2. In order to specify the basis functions, the degree was specified to be 2, in order to provide adequate smoothness, and one interior knot was placed at the median of the observed finite nonzero interval end points, with boundary knots being placed at the minimum and maximum of the same. The EM algorithm derived in Section 3.4 was used to complete model fitting for M1-M4. The starting value for all implementations was set to be $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\beta}^{(0)'}, \boldsymbol{\gamma}^{(0)'}, \alpha^{(0)}) = (\mathbf{0}_2', \mathbf{1}_k', 0.1)$, where $\mathbf{0}_k(\mathbf{1}_k)$ is a $(k \times 1)$-dimensional vector of zeros (ones). Convergence was declared when the maximum absolute difference between successive parameter updates was less than the specified tolerance of 1×10^{-5}.

Table 3.1 summarizes the estimates of the regression coefficients and the baseline instantaneous failure probability for all considered simulation configurations and models, when the observation times were drawn from an exponential distribution. This summary includes the empirical bias, the sample standard deviation of the 500 point estimates, the average of the 500 standard error estimates, and the empirical coverage probabilities associated with 95% Wald confidence intervals. Table 3.2 provides the analogous results for the case in which the observation times are sampled from a discrete uniform distribution. From these results, one will first note that across all considered simulation settings the proposed approach performs very well for M4 and the correct parametric model (i.e., M1 when $\Lambda_0(t) = \log(t+1)/\log(11)$ and M2 when $\Lambda_0(t) = 0.1t$); i.e., the parameter estimates exhibit very little bias, the sample standard deviation of the 500 point estimates are in agreement with the average of the standard error estimates, and the empirical coverage probabilities are at their nominal level. In summary, these findings tend to suggest that the proposed methodology can be used to reliably estimate the covariate effects, the instantaneous failure probability, and quantify

the uncertainty in the same. Additionally, these findings generally continue to persist for the case in which the parametric model is over specified (e.g., M3 when $\Lambda_0(t) = 0.1t$), with the resulting estimates in some cases exhibiting a slightly larger bias and a bit more variability, as one would expect. Further, from these results one will also note that when the parametric model is misspecified (e.g., M2 and M3 when $\Lambda_0(t) = \log(t+1)/\log(11)$) the estimates tend to exhibit more bias and less reliable inference, which is expected under the misspecification of the cumulative baseline hazard function. Finally, the estimates obtained under M4 (i.e., the model which makes use of the monotone splines) exhibit little if any difference when compared to the estimates resulting from the correct parametric model. In totality, from these findings, it is conjectured that the approach which makes use of the spline representation to approximate the unknown cumulative baseline hazard functions (i.e., M4) would likely be preferable, since it avoids the potential of model misspecification and it obtains estimators of the unknown parameters, as well as their standard errors, that are roughly equivalent to those estimators obtained under the true parametric model, the form of which is generally not known.

True $\Lambda_0(t) = \log(t+1)/\log(11)$																
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	-0.03	0.15	0.15	0.94	-0.06	0.17	0.15	0.89	-0.06	0.17	0.15	0.90	-0.04	0.16	0.16	0.93
$\beta_2 = -0.5$	-0.01	0.29	0.28	0.94	-0.05	0.33	0.27	0.89	-0.05	0.33	0.28	0.90	-0.02	0.30	0.29	0.93
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.95
$\beta_1 = -0.5$	-0.03	0.15	0.14	0.94	-0.07	0.17	0.14	0.87	-0.07	0.17	0.14	0.88	-0.04	0.15	0.15	0.94
$\beta_2 = 0.5$	0.01	0.26	0.26	0.94	0.06	0.30	0.25	0.89	0.06	0.30	0.26	0.89	0.03	0.27	0.27	0.94
$p = 0.3$	0.00	0.06	0.06	0.94	-0.01	0.06	0.05	0.92	-0.01	0.06	0.06	0.93	-0.01	0.06	0.06	0.94
$\beta_1 = 0.5$	0.01	0.15	0.15	0.95	0.05	0.16	0.15	0.91	0.05	0.16	0.15	0.92	0.02	0.15	0.15	0.95
$\beta_2 = -0.5$	0.01	0.28	0.28	0.94	-0.03	0.32	0.27	0.90	-0.03	0.32	0.27	0.91	0.01	0.29	0.29	0.94
$p = 0.3$	0.00	0.05	0.06	0.97	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96	0.00	0.05	0.06	0.96
$\beta_1 = 0.5$	0.02	0.14	0.14	0.95	0.07	0.15	0.14	0.89	0.07	0.15	0.14	0.90	0.03	0.14	0.15	0.95
$\beta_2 = 0.5$	0.02	0.28	0.26	0.94	0.06	0.32	0.25	0.86	0.06	0.32	0.26	0.87	0.03	0.28	0.27	0.93
$p = 0.3$	0.00	0.06	0.06	0.93	-0.01	0.06	0.05	0.90	-0.01	0.06	0.06	0.91	0.00	0.06	0.06	0.93
True $\Lambda_0(t) = 0.1t$																
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	0.01	0.15	0.16	0.96	-0.03	0.16	0.16	0.94	-0.05	0.17	0.16	0.94	-0.04	0.16	0.17	0.94
$\beta_2 = -0.5$	0.04	0.28	0.29	0.95	0.00	0.30	0.30	0.95	-0.03	0.32	0.30	0.94	-0.02	0.32	0.31	0.95
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95
$\beta_1 = -0.5$	0.02	0.15	0.15	0.94	-0.02	0.15	0.15	0.94	-0.04	0.16	0.15	0.93	-0.04	0.16	0.16	0.94
$\beta_2 = 0.5$	-0.02	0.26	0.27	0.96	0.02	0.28	0.27	0.93	0.03	0.29	0.28	0.93	0.03	0.29	0.29	0.95
$p = 0.3$	0.01	0.06	0.06	0.94	0.00	0.06	0.06	0.94	-0.01	0.06	0.06	0.94	-0.01	0.06	0.06	0.94
$\beta_1 = 0.5$	-0.02	0.14	0.16	0.96	0.02	0.15	0.16	0.96	0.03	0.16	0.16	0.95	0.03	0.16	0.17	0.95
$\beta_2 = -0.5$	0.06	0.29	0.29	0.94	0.02	0.30	0.29	0.93	0.01	0.31	0.30	0.93	0.01	0.31	0.31	0.94
$p = 0.3$	0.00	0.05	0.06	0.97	0.00	0.05	0.06	0.96	0.00	0.05	0.06	0.96	0.00	0.06	0.06	0.97
$\beta_1 = 0.5$	-0.02	0.14	0.15	0.95	0.02	0.15	0.15	0.95	0.04	0.15	0.15	0.95	0.04	0.15	0.16	0.95
$\beta_2 = 0.5$	-0.02	0.27	0.27	0.94	0.02	0.29	0.27	0.94	0.04	0.30	0.28	0.93	0.04	0.30	0.29	0.93
$p = 0.3$	0.01	0.06	0.06	0.96	0.00	0.06	0.06	0.95	-0.01	0.06	0.06	0.94	-0.01	0.06	0.06	0.95

Table 3.1: Summary of regression coefficient estimates and baseline instantaneous failure probability obtained from M1-M4 across all simulation settings, when the observation times were sampled from an exponential distribution. This include the average of the 500 point estimates minus the true value (Bias), the sample standard deviation of the 500 point estimates (SD), the average of the estimated standard errors (ESE), and empirical coverage probabilities associated with 95% Wald confidence intervals (CP95). Note, when $\Lambda_0(t) = \log(t+1)/\log(11)$ then M1 is the true parametric model and when $\Lambda_0(t) = 0.1t$ then M2 is the true parametric model.

True $\Lambda_0(t) = \log(t+1)/\log(11)$																
	M1(True)				M2(Misspecified)				M3(Misspecified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	-0.02	0.14	0.15	0.95	-0.03	0.15	0.15	0.92	-0.03	0.15	0.15	0.93	-0.03	0.15	0.15	0.94
$\beta_2 = -0.5$	0.00	0.27	0.27	0.96	-0.02	0.29	0.27	0.94	-0.02	0.29	0.27	0.94	-0.02	0.28	0.28	0.96
$p = 0.3$	-0.01	0.06	0.06	0.95	-0.01	0.06	0.06	0.93	-0.01	0.06	0.06	0.94	-0.01	0.06	0.06	0.94
$\beta_1 = -0.5$	-0.03	0.14	0.14	0.94	-0.05	0.15	0.14	0.92	-0.05	0.15	0.14	0.92	-0.04	0.15	0.14	0.94
$\beta_2 = 0.5$	0.00	0.26	0.25	0.94	0.03	0.28	0.25	0.92	0.03	0.28	0.25	0.93	0.02	0.27	0.26	0.93
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.95
$\beta_1 = 0.5$	0.02	0.16	0.15	0.94	0.03	0.16	0.15	0.94	0.03	0.16	0.15	0.94	0.03	0.16	0.15	0.94
$\beta_2 = -0.5$	0.00	0.28	0.27	0.95	-0.01	0.29	0.27	0.93	-0.01	0.29	0.27	0.93	-0.01	0.29	0.28	0.94
$p = 0.3$	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95
$\beta_1 = 0.5$	0.01	0.14	0.14	0.95	0.04	0.15	0.14	0.92	0.04	0.15	0.14	0.93	0.03	0.14	0.14	0.94
$\beta_2 = 0.5$	0.03	0.27	0.25	0.94	0.05	0.29	0.25	0.91	0.05	0.29	0.25	0.91	0.04	0.28	0.26	0.92
$p = 0.3$	0.00	0.06	0.06	0.95	-0.01	0.06	0.05	0.94	-0.01	0.06	0.06	0.94	0.00	0.06	0.06	0.94
True $\Lambda_0(t) = 0.1t$																
	M1(Misspecified)				M2(True)				M3(Over specified)				M4(Spline)			
Parameter	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95	Bias	SD	ESE	CP95
$\beta_1 = -0.5$	0.00	0.15	0.15	0.95	-0.02	0.15	0.15	0.93	-0.03	0.16	0.16	0.93	-0.03	0.16	0.16	0.94
$\beta_2 = -0.5$	0.02	0.27	0.28	0.96	0.00	0.28	0.28	0.96	-0.01	0.29	0.29	0.96	-0.01	0.29	0.29	0.96
$p = 0.3$	-0.01	0.06	0.06	0.95	-0.01	0.06	0.06	0.95	-0.01	0.06	0.06	0.94	-0.01	0.06	0.06	0.95
$\beta_1 = -0.5$	0.00	0.14	0.14	0.95	-0.02	0.14	0.14	0.94	-0.03	0.15	0.15	0.94	-0.03	0.15	0.15	0.93
$\beta_2 = 0.5$	-0.03	0.25	0.26	0.95	0.00	0.27	0.26	0.94	0.01	0.27	0.27	0.94	0.01	0.27	0.27	0.94
$p = 0.3$	0.01	0.06	0.06	0.96	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95
$\beta_1 = 0.5$	0.00	0.15	0.15	0.94	0.02	0.16	0.15	0.94	0.02	0.16	0.16	0.94	0.02	0.16	0.16	0.94
$\beta_2 = -0.5$	0.01	0.28	0.28	0.95	-0.01	0.29	0.28	0.95	-0.01	0.30	0.28	0.94	-0.01	0.30	0.29	0.94
$p = 0.3$	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.96
$\beta_1 = 0.5$	-0.02	0.13	0.14	0.96	0.01	0.14	0.14	0.95	0.02	0.14	0.15	0.95	0.02	0.14	0.15	0.96
$\beta_2 = 0.5$	0.00	0.26	0.26	0.96	0.03	0.27	0.26	0.95	0.04	0.28	0.27	0.95	0.04	0.28	0.27	0.95
$p = 0.3$	0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.95	0.00	0.06	0.06	0.96

Table 3.2: Summary of regression coefficient estimates and baseline instantaneous failure probability obtained from M1-M4 across all simulation settings, when the observation times were sampled from discrete uniform distribution. This include the average of the 500 point estimates minus the true value (Bias), the sample standard deviation of the 500 point estimates (SD), the average of the estimated standard errors (ESE), and empirical coverage probabilities associated with 95% Wald confidence intervals (CP95). Note, when $\Lambda_0(t) = \log(t+1)/\log(11)$ then M1 is the true parametric model and when $\Lambda_0(t) = 0.1t$ then M2 is the true parametric model.

Figure 3.1 summarizes the estimates of the baseline survival function (i.e., $S_0(t) = S(t|\mathbf{x} = \mathbf{0}_r)$) obtained from M1-M4 across all considered regression parameter configurations when $\Lambda_0(t) = \log(t+1)/\log(11)$ and the observation times were sampled from the exponential distribution. Figures 3.2-3.4 summarizes the analogous results for the other simulation configurations. In particular, these figures present the true baseline survival functions, the average of the point-wise estimates, and the point-wise 2.5th and 97.5th percentiles of the estimates. These figures reinforce the main findings discussed above; i.e., M4 and the correctly specified parametric model again provide reliable estimates of the baseline survival function, and hence the cumulative baseline hazard function, across all simulation configurations. Similarly the over specified model also provides reliable estimates, but the same can not be said for the misspecified models. It is worthwhile to point out that the baseline survival curves do not extend to unity as time goes toward the origin, this is due to the fact that the baseline instantaneous failure probability is $p = 0.3$. Again, these findings support the recommendation that the spline based representation of the cumulative baseline hazard function should be adopted in lieu of a parametric model, thus obviating the possible ramifications associated with misspecification.

In summary, this simulation study illustrates that the proposed methodology can be used to analyze current status data which is subject to instantaneous failures, and moreover that the monotone spline approach discussed in Section 3.2 should be adopted for approximating the unknown cumulative baseline hazard function. A few additional details about the numerics of the approach follow. First, the average time required to complete model fitting was approximately one second, supporting the claim that the proposed approach is computationally efficient. Lastly, for reasons of complete transparency, for a single data set, among 8000, the OPG estimator under M4 resulted in a singular matrix, which prevented standard error estimation, and this issue was eaqsily resolved by slightly shifting the placement of the interior knot.

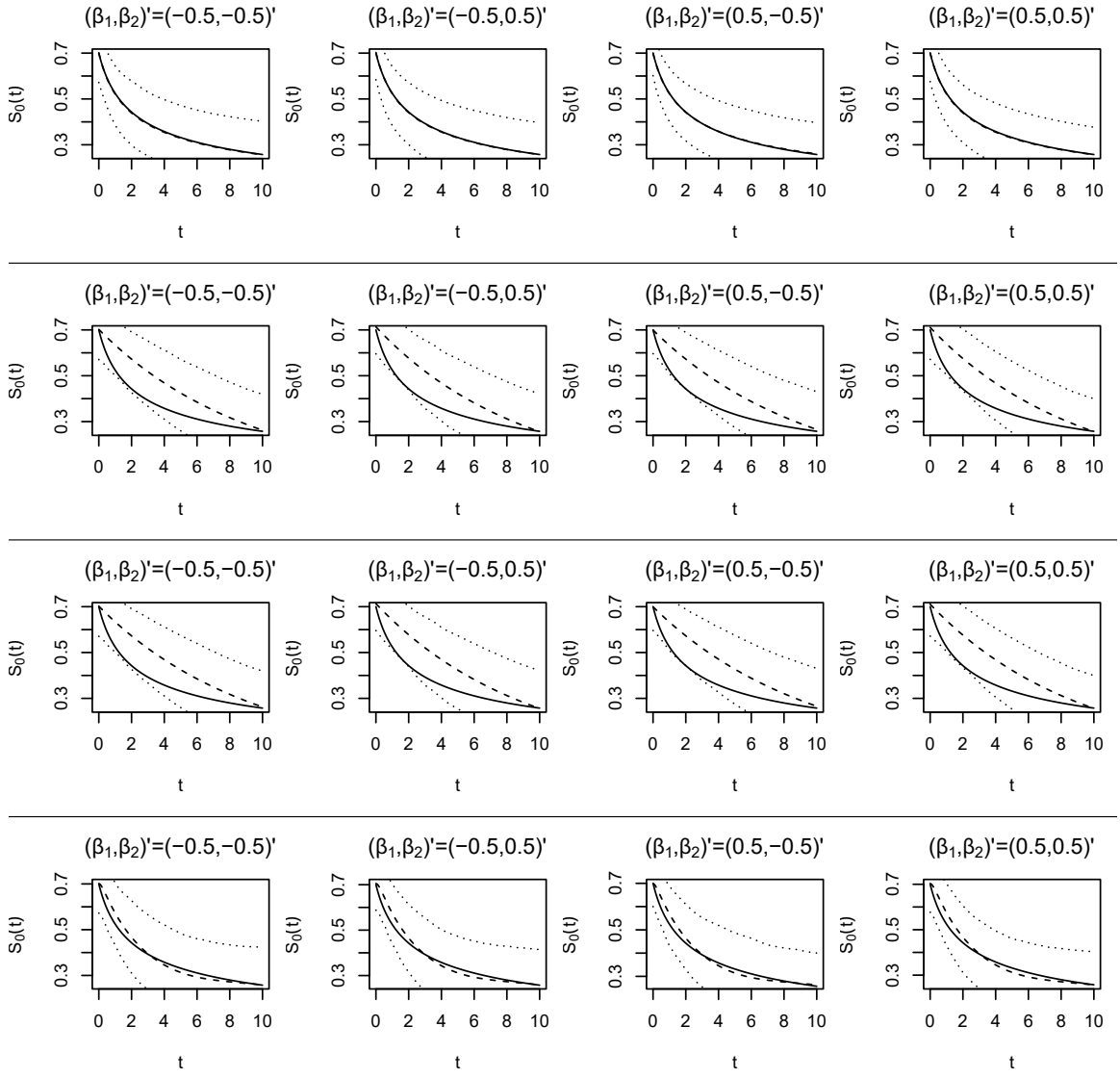

Figure 3.1: Simulation results summarizing the estimates of the baseline survival function obtained by the proposed approach under M1 (first row), M2 (second row), M3 (third row), and M4 (fourth row) when $\Lambda_0(t) = \log(t+1)/\log(11)$ and the observation times were drawn from an exponential distribution. The solid line provides the true value, the dashed line represents the average estimated value, and the dotted lines indicate the 2.5% and 97.5% quantiles, of the point-wise estimates. Note, M1 is the true parametric model in this setting.

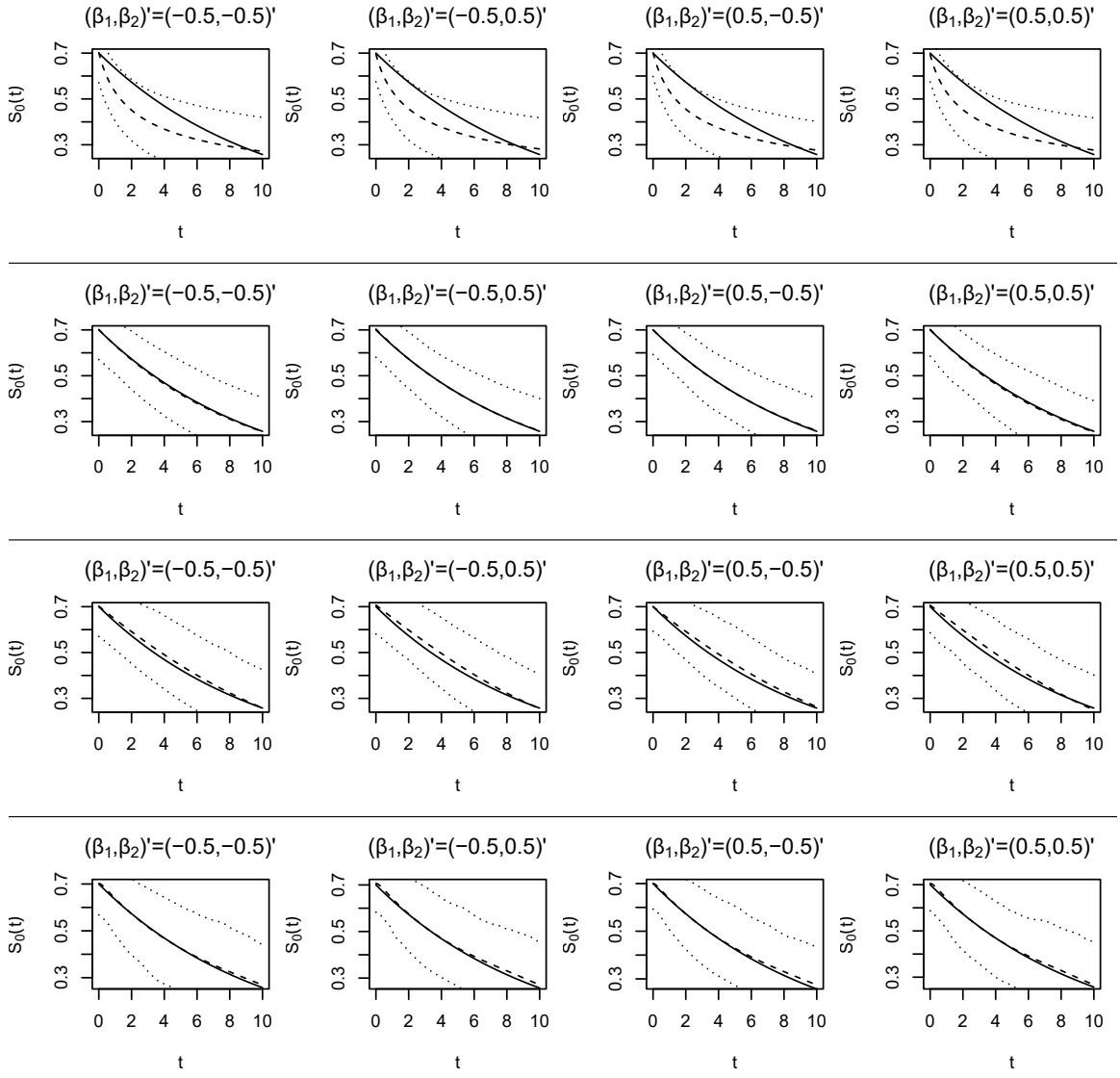

Figure 3.2: Simulation results summarizing the estimates of the baseline survival function obtained by the proposed approach under M1 (first row), M2 (second row), M3 (third row), and M4 (fourth row) when $\Lambda_0(t) = 0.1t$ and the observation times were drawn from an exponential distribution. The solid line provides the true value, the dashed line represents the average estimated value, and the dotted lines indicate the 2.5% and 97.5% quantiles, of the point-wise estimates. Note, M2 is the true parametric model in this setting.

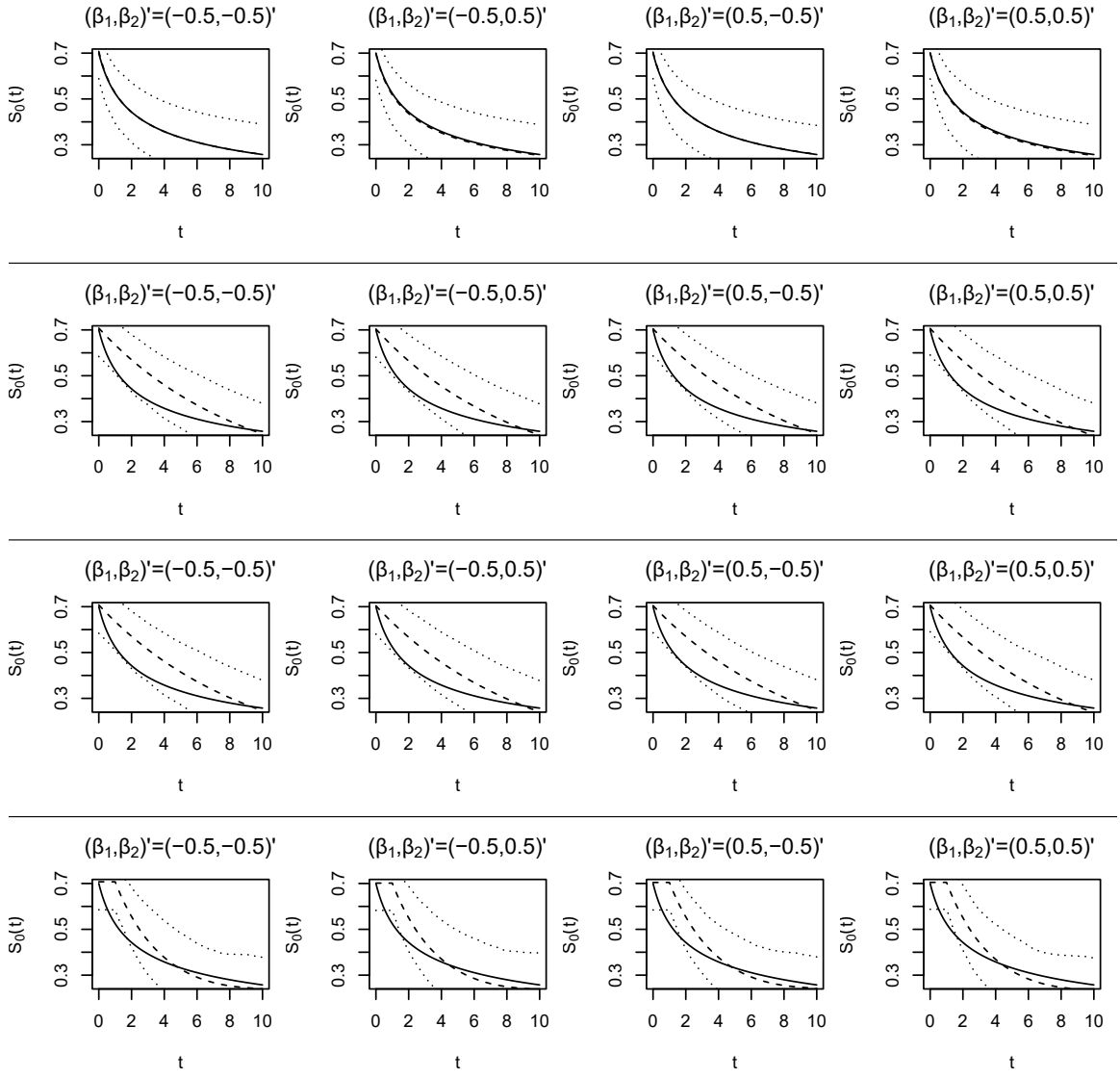

Figure 3.3: Simulation results summarizing the estimates of the baseline survival function obtained by the proposed approach under M1 (first row), M2 (second row), M3 (third row), and M4 (fourth row) when $\Lambda_0(t) = \log(t+1)/\log(11)$ and the observation times were drawn from an discrete uniform distribution over [1, 17]. The solid line provides the true value, the dashed line represents the average estimated value, and the dotted lines indicate the 2.5% and 97.5% quantiles, of the point-wise estimates. Note, M1 is the true parametric model in this setting.

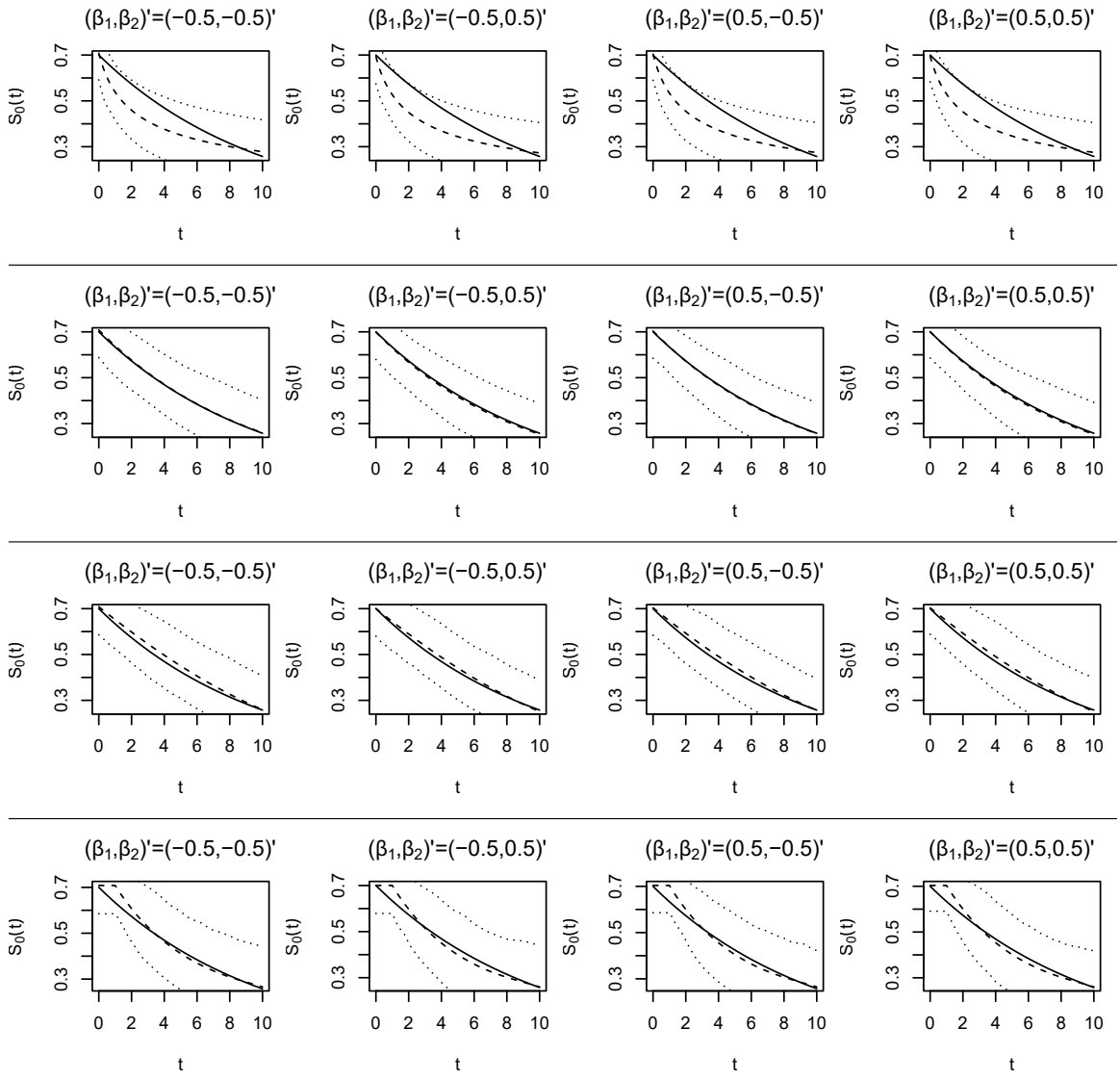

Figure 3.4: Simulation results summarizing the estimates of the baseline survival function obtained by the proposed approach under M1 (first row), M2 (second row), M3 (third row), and M4 (fourth row) when $\Lambda_0(t) = 0.1t$ and the observation times were drawn from an discrete uniform distribution over [1, 17]. The solid line provides the true value, the dashed line represents the average estimated value, and the dotted lines indicate the 2.5% and 97.5% quantiles, of the point-wise estimates. Note, M2 is the true parametric model in this setting.

3.5 Discussion

This work is aimed at developing, assessing, and validating a methodology which can be used to analyze the data resulting from the clinical trial described in Section 2 once it is complete. To this end, this work proposed a new model, under the PH assumption, which can be used to analyze interval-censored data subject to instantaneous failures. Through the model development, two techniques for approximating the unknown cumulative baseline hazard function are illustrated. To complete model fitting, a novel EM algorithm is developed through a two-stage data augmentation process. The resulting algorithm is easy to implement and is computationally efficient. These features are attributable to the fact that the carefully structured data augmentation steps lead to closed-form expressions for all necessary quantities in the E-step of the algorithm. Moreover, in the M-step the regression coefficients are updated through solving a low-dimensional system of equations, while all other unknown parameters are updated in closed-form. The finite sample performance of the proposed approach was exhibited through an extensive numerical study. This study suggests that the use of the monotone spline representation of the cumulative baseline hazard function would in general be preferable, in order to circumvent the possibility of model misspecification. To further disseminate this work, code, written in R, has been prepared and is available upon request from the corresponding author.

3.6 Acknowledgments

This research was partially supported by National Institutes of Health grants AI121351 and UL1 TR001111.

CHAPTER 4: STUDY DESIGN WITH STAGGERED SAMPLING TIMES FOR EVALUATING SUSTAINED UNRESPONSIVENESS TO PEANUT SLIT

In this work, we delineate an altered study design of a pre-existing clinical trial that is currently being implemented at the Department of Pediatrics at the University of North Carolina at Chapel Hill. The purpose of the ongoing investigation of the desensitized pediatric cohort is to address effectiveness of sublingual immunotherapy (SLIT) in achieving sustained unresponsiveness (SU) through the use of repeated Double Blind Placebo Control Food Challenge (DBPCFC) tests. SU is defined as the capacity to maintain desensitization to the food allergen following cessation of therapy. With scarce published literature defining and characterizing SU, the length of time off-therapy that would represent clinically meaningful benefit remains undefined. In this study, we use the new design features to assess time to loss of SU as an important efficacy endpoint that, to our knowledge, no prior study has investigated. Our work has two-fold objectives: first, to propose and discuss aspects of the altered design that would allow us to study SU; second, to explore and apply methodology to evaluate the time to loss of SU and risk factors in context of the data originating from the proposed study design. The salient feature of the new design is the allocation scheme of study subjects to staggered sampling time-points when a subsequent DBPCFC is administered. Due to a fixed sequence of increasing allergen doses administered in the challenge test, a patient's true threshold at both occasions is interval-censored. Further, due to the timing of the subsequent DBPCFC, the time to loss of SU is either left- or right-censored. Additionally, some participants at study entry do not pass the DBPCFC, leading to what can be construed as an instantaneous failure. Thus, for the purposes of analyzing data arising from this clinical trial, we also examine a recently proposed mixture model, which was specifically designed for the analysis of interval-censored data subject to instantaneous failures. In particular, through in-depth numerical studies we examine the performance and power of this new methodology to inform the effect of risk factors on the hazard of loss of SU.

4.1 Introduction

The prevalence of peanut allergy and the subsequent fatal reactions attributed to this allergen has continued to rise over the past decade, particularly in the United States pediatric population [37; 71]. In general, allergic reactions to peanuts can present with minor symptoms to include itchiness, urticaria, swelling, eczema, sneezing, and abdominal pain with more severe reactions leading to asthma, anaphylaxis, and even cardiac arrest. Although the current standard of care is a strict avoidance of peanut and ready-access to self-injectable epinephrine, implementation of an avoidance diet is complex. This leaves the majority of patients with a life-long risk of anaphylaxis and death due to accidental ingestions. Therapeutic interventions that provide everlasting protection from inadvertent allergen ingestion are, thus, needed.

Recently, the research surrounding the treatment of food allergies has shifted to consider different types of immunotherapy, through exposing patients to small doses of the allergen. The primary difference between different immunotherapy treatments involve the delivery mechanism. For example, subcutaneous immunotherapy (SCIT) exposes patients to the allergen through injections, oral immunotherapy (OIT) exposes patients to the allergen through direct ingestion, and sublingual immunotherapy (SLIT) involves the patient holding allergen extract under the tongue. For the purposes of desensitizing patients to food allergies, SCIT has shown little promise with a high rate of systematic reactions, while OIT and SLIT have shown promise, but further investigations need to be conducted. On other hand, SLIT, when compared to OIT, is often favored because it is thought to be safer, results in fewer systemic reactions, and has an easier mode of administration. However, with regard to the effectiveness of OIT or SLIT, almost all OIT and SLIT studies have demonstrated clinical desensitization; i.e., the patients experience an increase in reaction threshold while receiving the study drug [76], few studies have been successful in assessing tolerance or sustained unresponsiveness (SU). Here, SU refers to a non-reactive state that persists after discontinuation of therapy but wanes after a period [11]; whereas, tolerance refers to a non-reactive state of the immune system that persists after discontinuation of the study drug.

Various studies of OIT and SLIT have looked at SU [83; 82; 12] but these studies have either been small or have failed to investigate the loss of SU in a systematic fashion within a single population. The failure, in part, is due to study designs that require desensitized subjects to take subsequent

challenges at a specific time which could be too early to demonstrate SU. Considering that longer study durations run the risk of high subject withdrawal rate, such study designs are less pragmatic and more expensive. With the FDA approval of food immunotherapy potentially on the horizon, the development of study designs targeted at assessing SU is imperative. In light of this fact, and through having access to an extensive longitudinal dataset that is being collected on an ongoing clinical investigation of peanut SLIT in pediatric population, we are presented with a unique opportunity to address issues pertaining to SU with an innovative study design and advanced statistical methodology.

The Peanut SLIT Study is a revised augmentation of a pre-existing phase-II, prospective, randomized, open label 66-month clinical study supported by the National Institutes of Allergy and Infectious Diseases (NIAID) and conducted in the Allergy, Immunology, Rheumatology and Infectious Diseases (AIRID) Division in the Department of Pediatrics at the University of North Carolina at Chapel Hill [26]. We note that opposite of most studies that focus on estimating treatment effect, the study has an objective to assess the effect of taking away therapy. During the first 48 months of this study, participants were exposed to a build-up/maintenance phase during which they received peanut SLIT therapy in an incremental fashion during the initial 6 months and then treatment was maintained at a constant level thereafter. At the conclusion of this phase, a Double Blind Placebo Control Food Challenge (DBPCFC) was administered. Following the 48th month DBPCFC, the original protocol randomized desensitized subjects in a blinded fashion (2:1) to a placebo or treatment group, with the treatment group continuing peanut SLIT for 6 months. To assess clinical SU, the placebo group would complete the trial with a 54th month DBPCFC. In contrast, the treatment group would continue open label SLIT for an additional 6 months to complete the trial with a 66th month DBPCFC. However, results from [83] that indicated only 50% possessed SU after a month of using the presumed more robust OIT treatment lead to suspicion that 6 months on SLIT would work. Recognizing that a 6 month time window to assess clinical SU was rather optimistic, arbitrary, less supported by prior research, and that the study could result in a negligible success rate, the protocol was revised with staggered sampling times beyond the dose maintenance period for administration of the final DBPCFC among desensitized subjects. This dissolved the need for a two-arm trial and shortened the length of the overall clinical trial. Since time to loss of SU is an important dimension that, to our knowledge, no prior study has investigated, the revised study offers a robust framework and a systematic approach to study this clinically valuable information.

The revised study design presents an interesting case of censored data. First, due to the nature of the food challenge, a subject's true, time-dependent thresholds are obscured but are known relative to the no- and lowest-observed adverse event dose levels of the administered DBPCFCs. Second, because subject's true threshold in the absence of therapy is assumed to be guided by a monotonically non-increasing latent process, the time to loss of SU to a targeted dose administered at staggered observation times is either left- or right-censored; Last but not least, at the 48th month DBPCFC, subjects that exhibit a threshold lower than MCRT or the targeted dose level are assumed to have lost tolerance for the targeted dose at the study initiation; i.e., in the context of evaluating time to loss of SU, these subjects experience what is commonly referred to as an instantaneous failure. Thus, in addition to outlining the key features of the aforementioned study design, an alternate yet complimentary focus of this work is to evaluate a new statistical methodology that can be used to analyze the data arising from this clinical trial.

The statistical methodology in question was developed and presented in [89], and is an adaptation of the methods developed in [86]. This approach makes use of a generalization of the usual proportional hazards (PH) model. In particular, this generalization considers the use of a mixture model which allows one to account for instantaneous failures as well as interval-censored data, of which current status data is a special case. Under this mixture model, the covariates provide for a multiplicative effect on the hazard of experiencing a failure at both time zero and thereafter, thus obeying the PH assumption. Further, the approach uses a flexible monotone spline representation to approximate the cumulative baseline hazard function in the PH model. To complete model fitting, [89] developed, through a data augmentation process involving latent Poisson variables, an expectation-maximization (EM) algorithm. Through numerical studies, this method has shown a great deal of promise with respect to providing reliable estimation and inference, when used to analyze interval-censored data subject to instantaneous failures. The goal of the numerical studies presented herein is to validate this methodology in context of the data emanating from the clinical trial upon completion. In particular, this study examines the power of this procedure to detect significant effects.

The remainder of this article is organized as follows. In Section 2, we present the novel study design, and discuss in depth the ongoing clinical trial. Section 3 briefly outlines the salient features of the chosen statistical methodology. Section 4 reports the results from two extensive simulation

studies which were aimed at evaluating the finite sample performance of the proposed method when used to analyze data arising from the proposed study design. Section 5 concludes with a summary discussion.

4.2 Study Design and Primary Analysis

4.2.1 The Study Design

The study consists of a screening/baseline visit, build-up phase (approximately, 20 weeks), maintenance phase (42 months) and lastly, SU phase (17 weeks). In total, three DBPCFCs are administered, one at the baseline for subject screening and the other two at the end of the latter two phases for assessing desensitization and loss of SU thresholds, respectively. Upon enrollment, 51 children between the ages of 1 to 11 years underwent an entry DBPCFC with 1000 mg of cumulative peanut protein (25 mg, 50 mg, 100 mg, 250 mg, 575 mg) to confirm the peanut allergy diagnosis and establish a baseline threshold level. Following a positive DBPCFC, each subject was required to begin peanut SLIT at a starting dose of 2500 mcg to build up in increments to 4000 mcg during the build-up phase and to maintain this dose daily over the maintenance phase. At the end of maintenance phase, i.e., after 48 months of peanut SLIT administration, subjects were required to undergo a second DBPCFC with 5000 mg of cumulative peanut protein to assess desensitization. Subjects who were not able to consume more than the MCRT without symptoms were not considered desensitized and were required to stop peanut SLIT and conclude the study. Before the beginning of the SU phase, a computer-generated block randomization schedule was prepared such that each subject had an equal probability of being randomized to one of the weekly challenges held in the initial 6 weeks, between 7-12 weeks and between 13-17 weeks. Based on this schedule, three qualified subjects were randomly assigned to one of the final DBPCFC administered each week during the study's 17-week SU phase to evaluate for the loss of SU effect. The final DBPCFC, which is similar in structure to that of the 48th month, marks study completion. Since without continued antigen exposure, changes to the immune system and a continued increase in threshold is not expected, no individual is administered a final challenge dose higher than the highest administered dose that he/she was desensitized to at the 48th month DBPCFC. At the primary investigator's clinical discretion, subjects are then recommended to transition to a daily peanut food equivalent to maintain the desensitized effect.

4.2.2 The DBPCFC Structure and Time boundaries

In contrast to the standard challenges designed in such studies with multiplicative increments of a dose [68; 78], the similarly structured 48th month and the final DBPCFCs are inspired by the Fibonacci series that serves as a good check of observable cumulative effect of doses. The DBPCFC structure is illustrated in Figure 4.1. The rationale for the starting dose of 100 mg is based on the findings of [41] that had the median tolerated dose of 1700 mg with a wide range, lowest being 85 mg after a year long treatment. After 3+ years of treatment, the study had nobody with symptoms at less than 500 mg. The challenge size is kept at 5000 mg of cumulative peanut protein, chosen mostly arbitrarily but with a goal of avoiding false negative challenges as described in [68]. This is also in the range of what might be eaten; e.g., this is roughly equivalent to 1 tablespoon of peanut butter (4000 mg of peanut protein) or 1 serving size of peanuts (7000 mg of peanut protein). Considering these dose constraints, the end challenge is listed as (100, 200, 500, 800, 1300, 2100) mg with the first Fibonacci dose of 300 mg, also the minimal clinically relevant threshold level, broken into two incremental doses of 100 and 200 mg. Thus, compared to the structure with dose increments in multiples, this DBPCFC not only allows an observable MCRT of 300 mg, but also maintains the total maximum challenge size of 5000 mg. It also allows more granularity above the 500 mg dose level where most subjects are expected to reach their threshold. Further, the gradual dose gradation of this structure mitigates risk at the higher doses, providing enhanced safety from adverse events.

Each DBPCFC administered to a subject results in threshold bounds: NOAEL, the highest administered cumulative dose that elicits no adverse event and LOAEL, the lowest administered cumulative dose that elicits an adverse event [77]. Let $D_i^{T_0} \in [N_i^{T_0}, L_i^{T_0})$ define the ith subject's true desensitization threshold achieved at the 48th month ($T_0 = 0$) and $D_i^{T_{1i}} \in [N_i^{T_{1i}}, L_i^{T_{1i}})$, the true tolerance threshold achieved at the subject's final DBPCFC randomized time, T_{1i}, with corresponding NOAEL and LOAEL observed bounds such that $D_{MCRT}^{T_0} < N_i^{T_{1i}} \leq N_i^{T_0}$. Considering a subject's tolerance to a targeted dose $D_T = d$ in the absence of immunotherapy to be a non-increasing process, the time to loss of SU, T_i^d, is well approximated by an interval censoring framework as illustrated in Figure 4.1. Depicting two scenarios: i) $N_i^{T_{1i}} = N_i^{T_0}$; ii) $N_i^{T_{1i}} < N_i^{T_0}$, the subject's threshold trajectories end in an exhaustive set of possible targeted dose thresholds positioned relative to the observed thresholds. The wider dashed lines represent the subject's true latent trajectory representing

each scenario. For example, when the true latent trajectory is originating at 2500 and ending in 900 i.e. when $N_i^{T_{1i}} < N_i^{T_0}$, a subject is expected to meet a targeted threshold, say $D_T \leq N_i^{T_{1i}} = 800$ at or beyond T_{1i}; when $L_i^{T_0} = 2900 > D_T \geq L_i^{T_{1i}} = 1600$ and $D_T < L_i^{T_0} = 2900$, it would be before T_{1i}. Similarly, when the true latent trajectory is originating at 4000 and ending at 3200, i.e., when $N_i^{T_{1i}} = N_i^{T_0}$, a subject is expected to meet a targeted threshold, $D_T \leq N_i^{T_{1i}} = 2900$ at or beyond T_{1i}; however, when $D_T \geq L_i^{T_{1i}} = 5000$, a subject is considered to have lost tolerance at time $T_0 = 0$ discounting a rare positive outcome in favor of conservative treatment effect and assuming at large that a subject's threshold cannot rise in the absence of therapy. More precisely, the time to loss of SU to a targeted dose, T_i^d, including all possible scenarios illustrated in Figure 4.1, is summarized using the following observables:

$$T_i^d \in \begin{cases} 0, & \text{if } N_i^{T_0} \leq D_{MCRT}^{T_0} \text{ or } (N_i^{T_1 i} = N_i^{T_0} \text{ and } d \geq L_i^{T_{1i}} = L_i^{T_0}) \\ & \quad \text{or } (N_i^{T_1 i} < N_i^{T_0} \text{ and } d \geq L_i^{T_0}), \\ [0, T_{1i}), & \text{if } N_i^{T_1 i} < N_i^{T_0} \text{ and } L_i^{T_{1i}} \leq d < L_i^{T_0}, \\ [T_{1i}, \inf), & \text{if } d \leq N_i^{T_1 i} \leq N_i^{T_0}, \\ (0, \inf], & \text{if } L_i^{T_{1i}} > d > N_i^{T_{1i}} \text{ and } N_i^{T_1 i} \leq N_i^{T_0}. \end{cases}$$

We note that $N_i^{T_{1i}} \leq N_i^{T_0}$ is always true under the study design construct. Further, a target dose threshold that lies within an observed interval and not at the boundary does not contribute to any information on time to failure. A subject's loss of SU to an arbitrary targeted dose level, $D_T (> D_{MCRT}^{T_0})$, is earliest claimed at T_{1i} if $N_i^{T_{1i}} \geq D_T$. The non-desensitized subjects with $N_i^{T_0} \leq D_{MCRT}^{T_0}$ and those that exhibit $L_i^{T_0} \leq D_T$ are assumed to have lost tolerance for D_T at T_0 and are not administered the final challenge. This results in current status (or case-1 interval-censored) data which is subject to instantaneous failures. Further, it is worthwhile to note that due to the study design the censoring times are non-informative since the subjects are administered final DBPCFCs at randomized observation times; i.e., the observation process is independent of the failure times of interest.

4.3 Model and methods

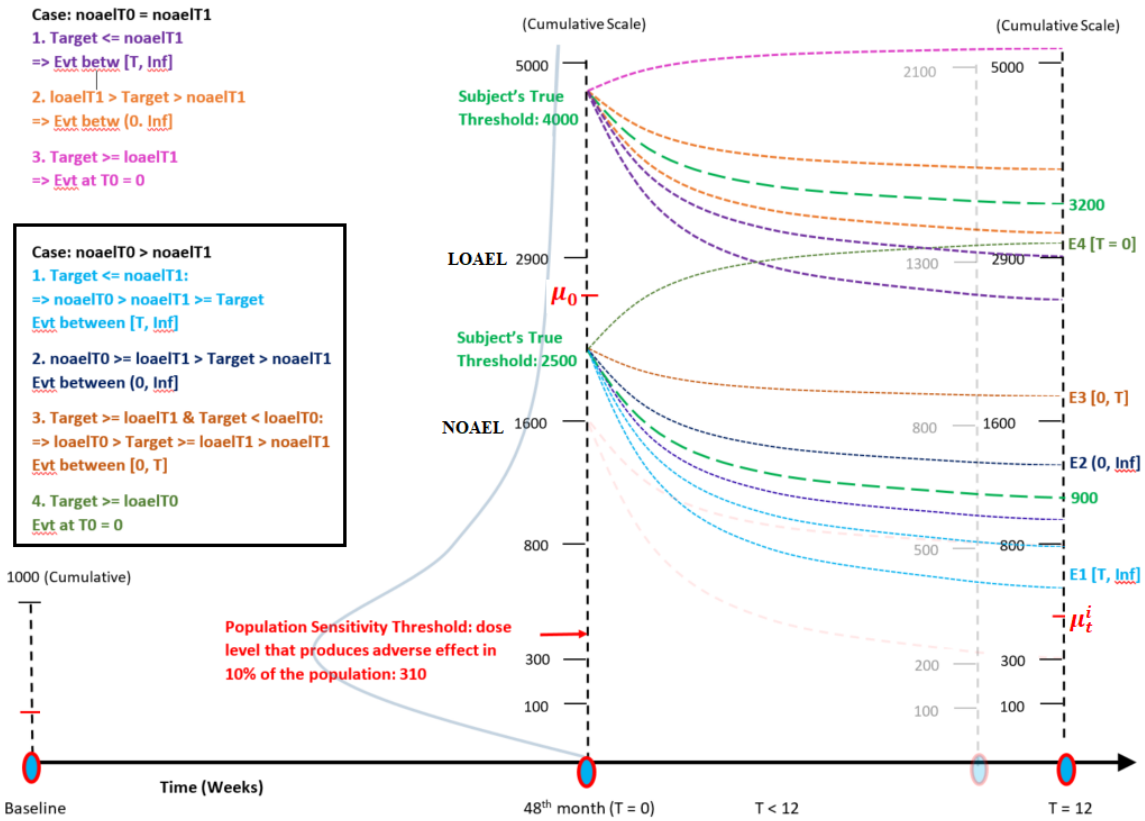

Figure 4.1: Peanut SLIT-TLC Study Design [26]: Subject's threshold trajectories ending in an exhaustive set of targeted thresholds relative to the observed thresholds for scenarios: i) when the NOAELs obtained at the 48th month and the final DBPCFCs are the same; ii) when the NOAEL obtained at the final DBPCFC is at least a level lower than that obtained at the 48th month. The broad dashed line represents a subject's true latent trajectory for either situation. All other trajectories end in targeted thresholds relative to the subject's final observed (NOAEL, LOAEL) and true latent threshold. For example, when the true latent final threshold is 900, a subject is expected to meet the targeted threshold of ≤ 800 at or beyond the time when the final DBPCFC is administered.

Herein, a brief discussion of the statistical methods is provided. Let $T^d \equiv T$ denote the survival time of interest for a fixed targeted dose d. The mixture model proposed in [89], provides the cumulative distribution function (cdf) of T as

$$H(t|\mathbf{x}) = \begin{cases} 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}, & \text{for } t = 0, \\ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} S(t|\mathbf{x}), & \text{for } t > 0\ , \end{cases} \tag{4.11}$$

where $\mathbf{x}$ is a $(r \times 1)$-dimensional vector of covariates, $\boldsymbol{\beta}$ is the corresponding vector of regression coefficients, $S(t|\mathbf{x}) = \exp\{-\Lambda_0(t)\exp(\mathbf{x}'\boldsymbol{\beta})\}$ is the survival function associated with the usual PH model, and $\Lambda_0(\cdot)$ is the cumulative baseline hazard function. A few comments are warranted. First, the covariates provide for a multiplicative effect on the hazard of experiencing a failure at both time zero and thereafter, thus obeying the PH assumption. Second, one will note that the mixture model allows for a probability mass at time 0 through the parameter α, this allows the model to account for the effect of instantaneous failures. Third, the baseline probability of experiencing an instantaneous failure in the absence of covariate effect is given by $p_0 \equiv P(T = 0|\mathbf{x} = \mathbf{0}_r) = 1 - S(T = 0|\mathbf{x} = \mathbf{0}) = 1 - \exp(-\alpha)$, where $\mathbf{0}_r$ is an $(r \times 1)$-dimensional vector of zeros. Thus, a subject with covariates $\mathbf{x}$ has a probability of experiencing an instantaneous failure given by

$$P(T = 0|\mathbf{x}) = 1 - [1 - p_0]^{e^{\mathbf{x}'\boldsymbol{\beta}}} = 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}.$$

Moreover, the survival function for individuals not experiencing an instantaneous failure is given by

$$P(T > t|\mathbf{x}) = P(T > 0|\mathbf{x})P(T > t|\mathbf{x}, t > 0) = e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} S(t|\mathbf{x}),$$

where $S(t|\mathbf{x})$ is as defined in (4.11).

The unknown parameters in the mixture model depicted in (4.11) involve the regression parameters $\boldsymbol{\beta}$, the baseline probability of experiencing an instantaneous failure p_0 (or equivalently α), and the cumulative baseline hazard function $\Lambda_0(\cdot)$. It is worthwhile to point out that for (4.11) to be a valid cdf $\Lambda_0(\cdot)$ must be a monotone increasing function, such that $\Lambda_0(0) = 0$. Moreover, this unknown function is best regarded as an infinite dimensional parameter. In order to reduce the dimensionality of the problem while allowing for adequate modeling flexibility, [89] suggest

modeling $\Lambda_0(\cdot)$ through the use of the monotone splines of [64]; i.e., specify

$$\Lambda_0(t) = \sum_{l=1}^{k} \gamma_l b_l(t), \tag{4.12}$$

where the $b_l(\cdot)$ are spline basis functions and the γ_l are unknown spline coefficient, for $l = 1, \ldots, k$. To assure that $\Lambda_0(\cdot)$ is nondecreasing, the γ_l should be nonnegative; i.e., $\gamma_l > 0$, for all $l = 1, \ldots, k$. Briefly, the spline basis functions are piecewise polynomial functions and are fully determined once a knot sequence and the degree are specified; for further details see [89; 86; 64].

4.3.1 Estimation and inference

For individuals not experiencing an instantaneous failure, let T_1 denote the time at which an individual is administered the final DBPCFC. This is when one learns whether or not the event of interest has occurred. In the study design outlined in Section 2, one will encounter three different scenarios; i.e., an instantaneous failure ($T = 0$), T is left-censored ($0 < T < T_1$), or T is right-censored ($T > T_1$). For notational convenience, let ψ be the indicator denoting time to failure is beyond time 0; i.e., $\psi = 1$, if T is not an instantaneous failure. Similarly, let δ be a censoring indicator taking value 1 if the observation is left-censored.

In order to derive the observed data likelihood, it is assumed that conditional on the covariates, the failure time for a subject is independent of the observational process. While common among the survival literature (see, e.g., [51; 94] and the references therein), this assumption is valid in this application by virtue of the study design; i.e., the study makes use of staggered randomized sampling times. The observed data collected on N subjects is given by $D = \{(T_{1i}, \mathbf{x}_i, \psi_i, \delta_i); i = 1, 2, \ldots, N\}$, which constitute N independent realizations of $(T_1, \mathbf{x}, \psi, \delta)$. Thus, under the aforementioned assumptions, the observed data likelihood can be expressed as

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} \left[F(T_{1i}|\mathbf{x}_i)^{\delta_i} \{1 - F(T_{1i}|\mathbf{x}_i)\}^{1-\delta_i} \right]^{\psi_i} \left\{ e^{-\alpha e^{\mathbf{x}_i'\boldsymbol{\beta}}} \right\}^{\psi_i} \left\{ 1 - e^{-\alpha e^{\mathbf{x}_i'\boldsymbol{\beta}}} \right\}^{1-\psi_i}, \tag{4.13}$$

where $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \alpha)'$ represents the set of unknown parameters which need to be estimated, with $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_k)'$. Since the observed data likelihood exists in closed-form, one could obtain the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$, as $\hat{\boldsymbol{\theta}} = \text{argmax}_{\boldsymbol{\theta}} L(\boldsymbol{\theta})$, through numerically maximizing the logarithm of (4.13). However, due to numerical instabilities of this process [87], it is generally

suggested that the EM algorithm developed and detailed in [89] be used for the purposes of identifying the MLE, which we denote herein as $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}', \hat{\boldsymbol{\gamma}}', \hat{\alpha})'$.

For uncertainty quantification, two variance estimators were considered and evaluated: the Huber sandwich estimator and the outer product of gradients (OPG) estimator. While both estimators in general seem to be viable, after extensive numerical studies (results not shown), the Huber sandwich estimator was found to be favorable in applications similar to the one described in Section 2 with no convergence issues for smaller sample sizes. In general, the Huber sandwich estimator is obtained as

$$\widehat{V}(\widehat{\boldsymbol{\theta}}) = -A^{-1}B(-A)^{-1},$$

where $A = \sum_{i=1}^{n} \partial^2 l_i(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^2|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$, $B = \sum_{i=1}^{n} \hat{g}_i(\widehat{\boldsymbol{\theta}})\hat{g}_i(\widehat{\boldsymbol{\theta}})'$, $\hat{g}_i(\widehat{\boldsymbol{\theta}}) = \partial l_i(\boldsymbol{\theta})/\partial\boldsymbol{\theta}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$, and $l_i(\boldsymbol{\theta})$ is the log-likelihood contribution of the ith individual. Using this estimator, one can conduct typical Wald type inference. In the subsequent section, it is demonstrated that the methods outlined in this section are capable of providing both reliable estimation and inference, with regard to analyzing data that emulates the data which is being collected from the ongoing clinical trial described in Section 2.

4.3.2 Numerical experiments

We perform the following simulation studies to investigate the finite sample performance of the proposed procedure within context of the data expected from the study described in section 2.

4.3.2.1 Simulation Study I

The true cumulative distribution function of the failure times was specified to be

$$H(t|\mathbf{x}) = \begin{cases} 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}}, & \text{for } t = 0, \\ 1 - e^{-\alpha e^{\mathbf{x}'\boldsymbol{\beta}}} S(t|\mathbf{x}), & \text{for } t > 0, \end{cases}$$

where $\mathbf{x} = (x_1, x_2)'$, $x_1 \sim \text{N}(0, 1)$, $x_2 \sim$ Bernoulli(0.5), $p_0 \in \{0.1, 0.2, 0.3\}$, and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, where β_1 and β_2 each take on values -0.5 and 0.5, thus leading to 4 regression parameter configurations. Recall, the relation between p_0 and α, which is given by $\alpha = -\log(1 - p_0)$. The true cumulative baseline hazard function was specified to be $\Lambda_0(t) = 0.1t$. Since the study is underway and the data is still being collected, the choice of an exponential distribution for event times rests on a reasonable assumption that a subject's baseline hazard of losing SU remains approximately constant. For each of the considered regression parameter configurations, this model was used to generate 500 independent

data sets each consisting of individual level data for N subjects, to include their failure times T, where $N \in \{50, 100, 200, 400\}$.

Observation times, T_1, as employed in the study, were generated using a block permutation schedule such that each subject had an equal probability of being randomized to one of the following time periods: initial 6 weeks, between 7-12 weeks and between 13-17 weeks. A subject was then assigned to a week sampled from the block they were randomized to. By comparing the observation times to the event times, it was determined whether or not the event time for each individual was left- or right-censored; i.e., the event time T is either less than or greater than T_1 leading to either left- or right-censoring, respectively.

The specification of the monotone spline representation outlined in Section 3 for the cumulative baseline hazard function proceeded to use basis functions of degree 2 and one interior knot placed at the median of observation times and boundary knots placed at the minimum (maximum) of the same. Then to complete the model fitting, the EM algorithm developed in [89] was used. The convergence of the algorithm was declared when the maximum absolute difference of all of the consecutive parameter updates were less that 1×10^{-5}. Further, for each data set the aforementioned Huber sandwich estimator was used to estimate the standard errors for the estimated regression coefficients.

Table 4.1 provides a summary of the estimated regression coefficients across all of the considered parameter configurations for varying sample sizes, $N \in \{50, 100, 200, 400\}$. This summary includes the sample mean, the mean standard error estimates, as well as sample standard deviation of the 500 estimates. Additionally, using the aforementioned Huber sandwich estimator, we conducted Wald type inference for the set of hypotheses $H_0 : \beta_j = 0$ versus $H_1 : \beta_j \neq 0$ for each data set and estimated power of this test to detect the effect at 0.05 significance level. From these results one will note that even for small sample sizes, the bias is generally small and the sample standard deviation of the 500 point estimates are generally in agreement with the mean standard error estimates suggesting that the Huber sandwich estimator performs reasonably well in estimating the asymptotic variances. As one might expect, the empirical power increases with the sample size. Further, from these results, it appears that the power to detect a significant effect associated with a continuous covariate dominates the power associated with the dichotomous covariate, which again is expected. Also, the negative and positive values of β_1 are expected to result in approximately similar power due

to symmetry of the continuous covariate; however, this symmetry does not hold for the dichotomous covariate. Figure 4.2 illustrates these findings for the case when $p_0 = 0.2$.

Figure 4.3 summarizes the estimated baseline (i.e., when $\mathbf{x} = \mathbf{0}_r$) survival functions for the 4 different configurations of the regression parameters when $N = 50$ and $p_0 = 0.2$. This figure presents the true baseline survival functions (solid lines), the average of the point-wise estimates (dashed lines), and the 2.5th and the 97.5th point-wise quantiles of the estimates (dotted lines). From these curves, one will note that the estimated baseline survival functions are very close to the true values. These results illustrate that the proposed methodology accurately estimates the baseline survival function which is roughly equivalent to accurately estimating the cumulative baseline hazard function. Moreover, this figure also illustrates that the proposed procedure can accurately estimate p_0, which is evidenced by the difference between 1 and the estimated baseline survival functions at time 0.

		N = 50				N = 100				N = 200				N = 400			
p0	**Parameter**	**Est**	**SE**	**SEE**	**Power**	**Est**	**SE**	**SEE**	**Power**	**Est**	**SE**	**SEE**	**Power**	**Est**	**SE**	**SEE**	**Power**
0.1	$\beta 1 = -0.5$	-0.59	0.27	0.30	0.612	-0.54	0.17	0.19	0.910	-0.52	0.11	0.12	0.996	-0.51	0.08	0.09	1.000
	$\beta 2 = -0.5$	-0.59	0.53	0.51	0.204	-0.55	0.32	0.33	0.386	-0.52	0.22	0.23	0.670	-0.51	0.15	0.15	0.916
	$\beta 1 = -0.5$	-0.59	0.25	0.28	0.692	-0.54	0.16	0.17	0.934	-0.52	0.11	0.12	0.998	-0.51	0.08	0.08	1.000
	$\beta 2 = 0.5$	0.56	0.50	0.50	0.232	0.52	0.29	0.30	0.418	0.51	0.20	0.21	0.718	0.51	0.14	0.15	0.956
	$\beta 1 = 0.5$	0.58	0.32	0.32	0.572	0.53	0.17	0.17	0.910	0.52	0.11	0.13	0.998	0.51	0.08	0.08	1.000
	$\beta 2 = -0.5$	-0.61	0.58	0.52	0.206	-0.56	0.32	0.32	0.424	-0.53	0.22	0.24	0.682	-0.50	0.15	0.16	0.916
	$\beta 1 = 0.5$	0.59	0.27	0.28	0.667	0.53	0.15	0.16	0.954	0.53	0.11	0.12	1.000	0.51	0.08	0.08	1.000
	$\beta 2 = 0.5$	0.56	0.69	0.51	0.213	0.51	0.29	0.30	0.428	0.51	0.20	0.22	0.720	0.52	0.14	0.15	0.948
0.2	$\beta 1 = -0.5$	-0.56	0.26	0.28	0.636	-0.53	0.16	0.17	0.928	-0.52	0.11	0.12	1.000	-0.51	0.08	0.08	1.000
	$\beta 2 = -0.5$	-0.57	0.53	0.47	0.222	-0.55	0.30	0.31	0.446	-0.51	0.21	0.22	0.710	-0.51	0.14	0.15	0.934
	$\beta 1 = -0.5$	-0.56	0.23	0.26	0.730	-0.54	0.15	0.16	0.960	-0.52	0.10	0.11	1.000	-0.51	0.07	0.07	1.000
	$\beta 2 = 0.5$	0.56	0.47	0.47	0.262	0.52	0.28	0.28	0.454	0.52	0.19	0.20	0.776	0.51	0.13	0.14	0.974
	$\beta 1 = 0.5$	0.56	0.31	0.27	0.612	0.53	0.18	0.16	0.928	0.52	0.11	0.12	1.000	0.51	0.08	0.08	1.000
	$\beta 2 = -0.5$	-0.59	0.63	0.46	0.228	-0.55	0.59	0.30	0.422	-0.53	0.21	0.22	0.724	-0.51	0.14	0.15	0.938
	$\beta 1 = 0.5$	0.58	0.24	0.26	0.709	0.53	0.15	0.16	0.962	0.52	0.10	0.11	1.000	0.51	0.07	0.07	1.000
	$\beta 2 = 0.5$	0.54	0.48	0.42	0.222	0.51	0.28	0.27	0.472	0.51	0.19	0.20	0.746	0.51	0.13	0.14	0.966
0.3	$\beta 1 = -0.5$	-0.56	0.25	0.25	0.660	-0.53	0.15	0.16	0.942	-0.52	0.10	0.11	1.000	-0.51	0.07	0.08	1.000
	$\beta 2 = -0.5$	-0.57	0.55	0.44	0.214	-0.55	0.29	0.29	0.476	-0.52	0.20	0.21	0.738	-0.51	0.14	0.14	0.944
	$\beta 1 = -0.5$	-0.56	0.31	0.24	0.733	-0.54	0.14	0.15	0.970	-0.52	0.10	0.11	1.000	-0.51	0.07	0.07	1.000
	$\beta 2 = 0.5$	0.55	1.02	0.42	0.275	0.51	0.26	0.27	0.498	0.51	0.18	0.19	0.790	0.51	0.13	0.13	0.986
	$\beta 1 = 0.5$	0.56	0.26	0.26	0.668	0.53	0.15	0.16	0.950	0.52	0.10	0.12	0.998	0.51	0.07	0.07	1.000
	$\beta 2 = -0.5$	-0.59	0.87	0.44	0.248	-0.55	0.29	0.29	0.448	-0.52	0.20	0.22	0.740	-0.51	0.14	0.15	0.956
	$\beta 1 = 0.5$	0.57	0.26	0.24	0.737	0.53	0.14	0.15	0.978	0.52	0.10	0.11	1.000	0.51	0.07	0.07	1.000
	$\beta 2 = 0.5$	0.54	0.56	0.41	0.251	0.51	0.26	0.26	0.490	0.51	0.18	0.19	0.790	0.51	0.13	0.14	0.980

Table 4.1: Simulation Study I: Summary of regression parameter estimates across all considered configurations. This include the average of the 500 point estimates (Est), the average of the estimated standard errors (SE), the standard deviation of the 500 point estimates (SD), and empirical power (Power) to detect the effect at 0.05 significance level.

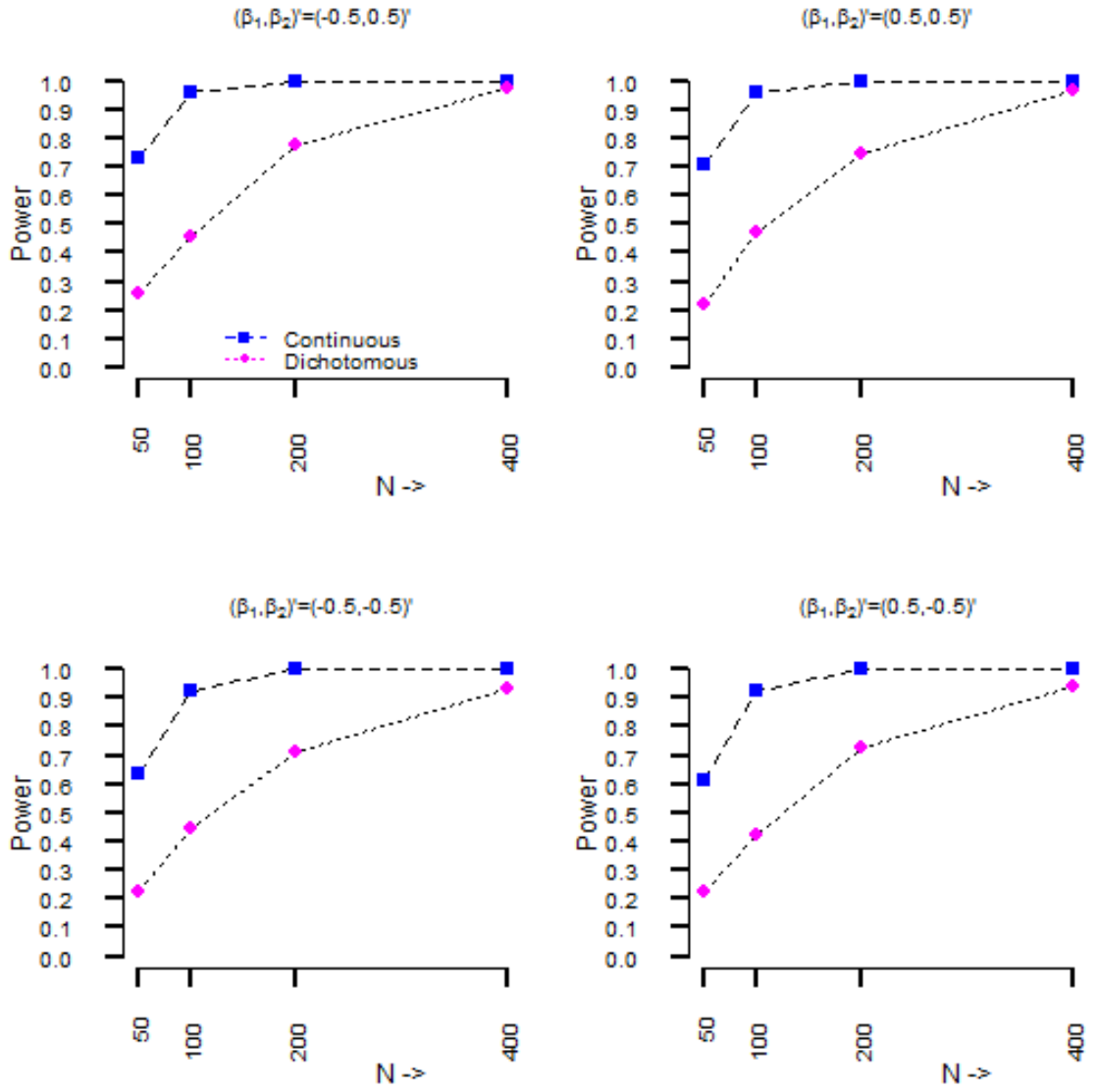

Figure 4.2: Simulation Study I: These figures provide empirical power curves, as a function of the sample size N, for the different configurations of the regression parameters. Note, the dashed and the dotted lines represent empirical power curves for the effects associated with the continuous and dichotomous covariates, respectively. The negative and positive values of β_1 are expected to result in approximately similar power due to symmetry of the continuous covariate; however, this symmetry does not hold for the dichotomous covariate.

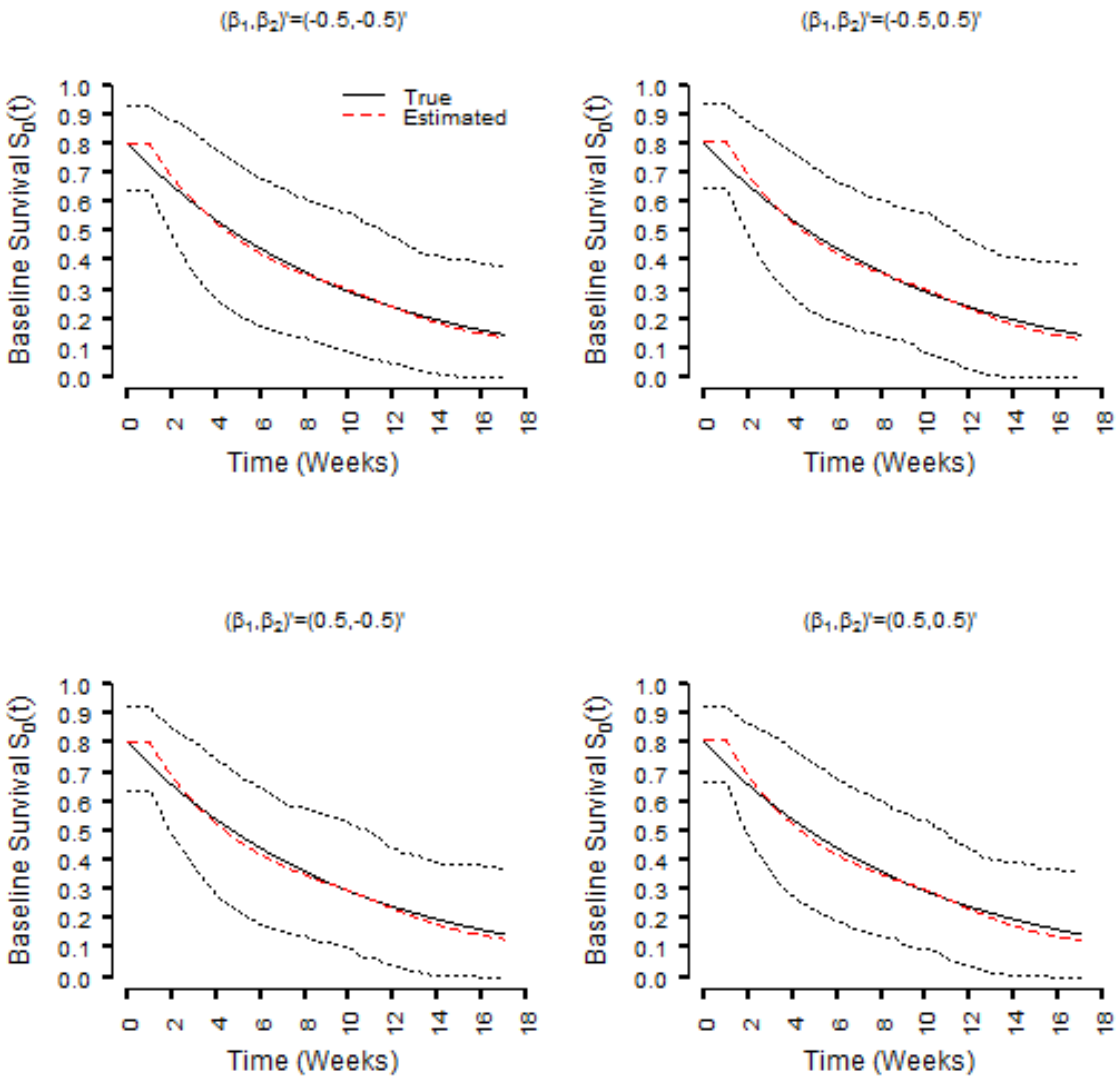

Figure 4.3: Simulation Study I: These figures depict the survival functions for the four different configurations of the regression parameters when $N = 50$ and $p_0 = 0.2$.Provided results include the true baseline survival function (solid lines), the average of the point-wise estimates (dashed lines), and the 2.5th and the 97.5th point-wise quantiles of the estimates (dotted lines).

4.3.2.2 Simulation Study II

This study was designed to emulate the data expected from the clinical trial outlined in Section 2. The data generating process in this study is identical to that described in the previous study, with a few minor alterations. In particular, since the clinical trial described in Section 2 has 51

participants, here we set $N = 50$. Further, this study sets the rate of instantaneous failure (i.e., the proportion of participants not successfully being desensitized by the 48th month DBPCFC), $p_0 = 0.2$, an average of those considered in the previous study. The goal of this study will be to examine the power of the procedure to detect significant effects ranging from a small to large signal level, thus we specify $\beta_1 \in \{0, 0.5, 1\}$ and $\beta_2 \in \{-1, -0.5, 0, 0.5, 1\}$ for the continuous and dichotomous covariate, respectively. Negative values of β_1 are not considered since the continuous covariate in the estimates is symmetric about zero. This symmetry does not hold for the dichotomous covariate and hence, negative values of β_2 are included. Table 4.2 provides the same summary offered in Table 4.1 but for the parameter estimates arising from this study, and these results reinforce the main findings discussed in the previous section. The primary interest in this study is to assess the power of the procedure with respect to the set of hypotheses given by $H_0 : \beta_j = 0$ versus $H_1 : \beta_j \neq 0$, as a function of the true effect size.

Figure 4.4 illustrates results associated with this aim. From this figure we see that the considered procedure will allow us to reliably detect important effects with a sample size of 50, as long as they are of a moderate size. It is worthwhile to point out that we do not assess the power of the procedure for the regression parameter associated with the continuous covariate, across negative effect sizes due to symmetry. Further, the size of the test is reasonably preserved. In summary, from the findings of these studies, we conjecture that the proposed methodology will be a reliable statistical approach which can be used to analyze data arising from the ongoing clinical trial outlined in Section 2.

Parameter	Est	SE	SEE	Power
$\beta 1 = 0.0$	0.00	0.27	0.28	0.098
$\beta 2 = -1.0$	-1.14	0.68	0.54	0.568
$\beta 1 = 0.0$	0.01	0.23	0.26	0.078
$\beta 2 = -0.5$	-0.57	0.52	0.46	0.214
$\beta 1 = 0.0$	0.00	0.22	0.24	0.074
$\beta 2 = 0.0$	-0.02	0.47	0.44	0.032
$\beta 1 = 0.0$	0.00	0.21	0.24	0.076
$\beta 2 = 0.5$	0.56	0.47	0.47	0.240
$\beta 1 = 0.0$	0.00	0.22	0.24	0.080
$\beta 2 = 1.0$	1.16	0.55	0.52	0.717
$\beta 1 = 0.5$	0.57	0.32	0.3	0.564
$\beta 2 = -1.0$	-1.17	0.92	0.54	0.598
$\beta 1 = 0.5$	0.56	0.31	0.27	0.612
$\beta 2 = -0.5$	-0.59	0.63	0.46	0.228
$\beta 1 = 0.5$	0.57	0.24	0.27	0.664
$\beta 2 = 0.0$	-0.04	0.46	0.45	0.052
$\beta 1 = 0.5$	0.58	0.24	0.26	0.708
$\beta 2 = 0.5$	0.54	0.48	0.42	0.220
$\beta 1 = 0.5$	0.57	0.24	0.25	0.691
$\beta 2 = 1.0$	1.12	0.52	0.48	0.727
$\beta 1 = 1.0$	1.15	0.36	0.36	0.940
$\beta 2 = -1.0$	-1.19	0.66	0.57	0.540
$\beta 1 = 1.0$	1.15	0.35	0.33	0.976
$\beta 2 = -0.5$	-0.61	0.74	0.47	0.200
$\beta 1 = 1.0$	1.14	0.34	0.34	0.964
$\beta 2 = 0.0$	-0.03	0.63	0.45	0.038
$\beta 1 = 1.0$	1.14	0.31	0.35	0.982
$\beta 2 = 0.5$	0.54	0.53	0.46	0.208
$\beta 1 = 1.0$	1.16	0.35	0.35	0.974
$\beta 2 = 1.0$	1.11	0.63	0.48	0.676

Table 4.2: Simulation Study II: Summary of the regression parameter estimates across all considered configurations. This includes the average of the 500 point estimates (Est), the average of the estimated standard errors (SE), the standard deviation of the 500 point estimates (SD), and empirical power (Power) to detect the effect at 0.05 significance level.

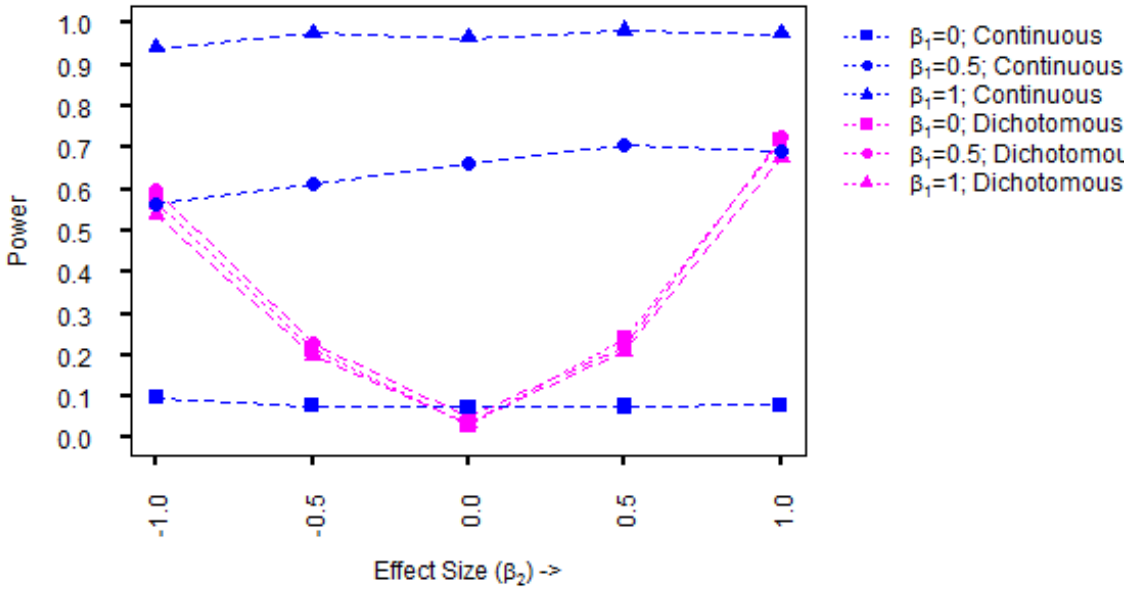

Figure 4.4: Simulation Study II: This figure provides empirical power curves, as a function of the effects sizes, for the different configurations of the regression parameters.

4.4 Discussion

The clinical design of this study emanates from a revised protocol of an ongoing 66-month SLIT study. The original study randomized desensitized subjects at the 48th month in a blinded fashion (2:1) to a placebo or treatment group, with the treatment group continuing peanut SLIT for 6 months. To assess clinical tolerance, the original protocol mandated that the placebo group complete the trial with a 54th month DBPCFC, while the treatment group completed the trial with a 66th month DBPCFC. Due to concerns about negligible success rates, the ongoing study was revised in hopes of effectively assessing clinical sustained unresponsiveness within patients after SLIT discontinuation. Since time to loss of tolerance is an important dimension that, to our knowledge, no prior study has investigated, the study incorporated amendments beyond 48 months of dose build-up and maintenance to provide a robust framework for studying this clinically valuable information.

The altered design also reflects an innovative structure of a food challenge inspired by a Fibonacci series that allows for more modest increments of allergen doses compared to a traditional design wherein dose is incremented in multiples. In contrast to the larger gaps of the traditional design between two consecutive higher dose levels, this challenge design helps mitigate subjects' risk by providing more opportunities of monitoring adverse events at higher doses of administered allergen. It is important that the key criterion of safety is satisfied, especially when action levels are set for regulatory purposes. Further, to ensure the identification of both NOAELs and LOAELs for all subjects, such a structure not only allows for a low-dose challenge to reduce the chances of left censoring, but also caps with high enough dose to limit right censoring. Also, due to narrower dose intervals, it allows for more information for analyses. Last but not least, it serves to observe the cumulative effect of doses which is more useful for the purposes of dose administration.

The presence of events at study entry (i.e., instantaneous failures) presents several challenges when one wishes to conduct a thorough statistical analysis. Traditionally, methods that conduct the regression analysis of censored survival data do not directly account for the effect of instantaneous failures. In fact, little to no guidance on how to directly address this scenario is available in the current literature. Typically, authors advocate introducing an arbitrary small constant, less than the smallest observation time, for all instantaneous failures. The study's smallest observation time is either shifted to this constant to have riskset and event-set include subjects with instantaneous failures

or the introduced constant is treated as the first observation time before the beginning of the study; i.e., these events after the introduction of the small constant are treated as left-censored observations. However, through numerical studies (results not shown) we have found that such strategies may be deleterious to both estimation and inference. In studies of food challenges that suffer from lack of a priori knowledge of threshold distribution, it is typical to have a minimal clinically relevant threshold or a targeted dose tolerance set at the investigators discretion. Such studies bear the risk of higher thresholds being set, which could lead to several failures at study initiation and hence could yield biased survival estimates, if traditional methods of handling instantaneous failures are employed. In light of these difficulties, we advocate the use of the methodology developed in [89], for the analysis of interval-censored data subject to instantaneous failures. Through extensive simulation studies, this new methodology has been shown to provide an accurate and efficient analysis in this context.

The simulation studies presented herein illustrate the finite sample performance of the methodology developed in [89] within the venue of the motivating clinical trial. These studies demonstrate that this methodology can accurately estimate all of the unknown parameters in the mixture model described in (4.11); i.e., the regression coefficients, the survival functions, and the baseline probabliltiy of experienceing an instantaneous failure. Further, the inferential techniques considered tend to preserve the specified size of the test, when the null is infact true, and exhibt sufficient power to detect departures from the null for both continuous and dichotomous covariates. Based on the numerical studies conducted herein, we believe the methodology developed in [89] will be capable of detecting clinically important effects in the data collected from the ongoing clinical trial, once it is complete. However, in our approach, we have not capitalized on the information available in the threshold distribution of DBPCFCs; instead, based on a subject's initial and final threshold intervals, we assume subjects have a non-increasing threshold trajectory to assess the period within which the targeted threshold is likely to be achieved. Consequently, this restricts our inference to only administered target dose thresholds rather than any arbitrary threshold in the distribution. However, we hope to build more sophisticated methods in sequel papers that would allow us to use the additional information on threshold trajectories.

In summary, we have illustrated a study design and accompanying method of analysis for interval-censored outcomes subject to instantaneous failures. By virtue of the study design, which has the final DBPCFCs at staggered sampling examination times, the observational process is

independent of the failure times, which is required by the methodology under study, as well as many previous proposals. The revised protocol has not only eliminated the need for a two-arm trial and hence, the associated administrative burden, but also reduced the overall length of the clinical trial by almost fifteen months. This feature significantly reduces the chance of non-compliance and drop-outs. Further, with a few changes in the study design and efficacy analysis plan, this revised protocol now will allow us to investigate new objectives of inquiry. We are hopeful that the knowledge shared here will help facilitate future studies in this therapeutic area.

4.5 Acknowledgments

This research was partially supported by National Institutes of Health grant AI121351 and UL1 TR001111.

CHAPTER 5: TOLERANCE DEGRADATION MODELING BASED ON TIME-DEPENDENT ORNSTEIN-UHLENBECK PROCESS

We introduce a time-dependent Ornstein-Uhlenbeck (OU) process for stochastic degradation of immunologic tolerance to peanut allergy. The model is a good choice due to its statistical properties on controllable mean, variance and correlation and its mean-reverting property that allows temporary correlated fluctuations from an overall degrading trend.

5.1 Ornstein-Uhlenbeck Process

5.1.1 General Process

Consider a one-dimensional structured differential equation (SDE) for an OU process:

$$dY_t = \big[AY_t + B\big]dt + \sigma dW_t$$

where dY_t represents change in tolerance threshold over a finite time interval, A, B, σ are constants, W_t is the standard Weiner motion.The process drift is explained by the parameter A and its diffusion by the parameter σ.

However, this model is restrictive and not flexible enough to be directly applied in degradation modeling. While it probably seems reasonable that the tolerance threshold satisfy the Markov property and have independent increments, it is not reasonable to assume that changes in thresholds are normally distributed; after all, we know that tolerance threshold can never fall below zero. Another issue of concern is a constant mean of the OU process. While there is no reason to presume changing variance, a constant mean disobeys an overall degrading trend. Further, in a biomedical experiment, a process is usually governed by a systematic trajectory that constitutes repeated measurements over time, but is often perturbed by individual differences among experimental units. To account for both system noise and random effects, stochastic differential mixed-effects models (SDMEM) naturally emanate. This extension, where system noise is modeled by including a diffusion term, allows for

simultaneous representation of randomness in the dynamics of the phenomena being considered and inter-individual variability.

We introduce three modifications to compensate the inflexibility of the earlier model: (i) to allow changes in thresholds to be log-normally distributed, we work with logarithm of threshold, $X_t = log(Y_t)$, to follow OU process given by the above SDE, which is equivalent to Y_t following a Geometric process: $dY_t = \big[A_g Y_t + B_g\big]dt + \sigma_g Y_t dW_t$ with difference in parametrization; (ii) start with introducing time dependent coefficients and later relax this assumption due to the sparse data emanating from the clinical trial that is less supportive of their estimation; (iii) allow systematic mean of the process to vary by individuals. As a consequence, a revised Ito's SDE incorporating random parameters is proposed. By using this methodology on repeated measurements from different subjects, it is not necessary to fit the individual data separately, but a single estimation procedure is used to fit the overall data simultaneously. The proposed model with additional assumptions that defines stochastic process for tolerance degradation modeling is considered in the following section.

5.1.2 SDMEM Framework

Consider a one-dimensional continuous process X_t evolving in M different experimental units (subjects) randomly chosen from a theoretical population. An SDMEM is defined as:

$$dX_t^i = A_t\big[X_t^i + \frac{B_t^i}{A_t}\big]dt + \sigma_t dW_t^i, \tag{5.14}$$

where X_t^i is logarithm of threshold degrading process at time $t \geq t_0^i$ in the i^{th} subject; $A_t = -\theta$, the drift (defines speed of reversion) and $\sigma_t = \sigma$, the diffusion parameter, are assumed to be non-negative constants similar for all subjects for model simplification; B_t^i is a time varying smooth function that varies across subjects; W_t^i is a standard Weiner motion. Additionally, we introduce a 1-dimensional subject-specific random effects parameter, $b^i \sim N(0, \sigma_b)$, that accounts for inherent individual differences across the entire temporal trajectory. Under the following notations:

$$\begin{aligned}
\alpha(t,s) &= -\int_s^t A_u du, \\
\beta^i(t,s) &= -\int_s^t B_u^i e^{\alpha(u,s)} du, \\
\gamma(t,s) &= \int_s^t C_u e^{2\alpha(u,s)} du
\end{aligned} \tag{5.15}$$

where $C_t = \frac{\sigma^2}{2} > 0$, an explicit form of X_t^i can be derived from (5.14):

$$X_t^i = e^{-\alpha(t,0)}\big[X_0^i + b^i - \beta^i(t,0) + \int_0^t \sigma e^{\alpha(u,0)} dW_u^i\big] \tag{5.16}$$

where X_0^i represents initial log threshold measured for an i^{th} subject at the 48^{th} month food challenge. Due to interval censored thresholds drawn from food challenge tests, the true threshold for a subject is unknown and assumed to be a random variable, $X_0^i \sim N(\mu_0, \sigma_0)$. The X_0^i, W_t^i and b^i are assumed independent of each other for all $1 \leq i \leq M$. We define the time varying smooth function $B_t^i = \theta\mu_t^i$, where $\mu_t^i = \mu(X_t^i, \gamma)$; $\gamma \in \Gamma \subseteq R^p$ is a p-dimensional fixed effects parameter (the same for entire population) and is defined as:

$$\begin{aligned} \mu_t^i &= \mu_\infty + (\mu_0 - \mu_\infty)e^{Z^i\Gamma - \delta t^i} \\ &= (\mu_0 - \mu_d) + \mu_d e^{Z^i\Gamma - \delta t^i} \end{aligned} \tag{5.17}$$

where $\mu_d = \mu_0 - \mu_\infty \geq 0$ and $E(e^{Z^i\Gamma})$ is constrained to equal 1 such that when $t^i = 0$, the conditional mean $\mu_{t=0}^i \equiv \mu_0^i = (\mu_0 - \mu_d) + \mu_d e^{Z^i\Gamma}$; when $t^i \to \infty, \mu_t^i \to \mu_\infty$, a constant. Here, Z^i is 1 x p dimensional matrix of subject's baseline covariates and the latest test results collected as of the 48th month challenge. Further, we re-parameterize X_0^i to incorporate random effect b_i, which is then expressed in terms of the population mean μ_0 as below:

$$X_0^i = (\mu_0 - \mu_d) + \mu_d e^{Z^i\Gamma} + b_i, \tag{5.18}$$

It should be emphasized that even though the OU process defines response over the entire time spectrum, $t^i > 0$, we observe only one instance of interval-censored threshold, $X_t^i \equiv X_2^i$ emanating from the final food challenge administered at randomized time-points for each individual. Thus, model (5.16) assumes that the process dynamics responsible for evolving X in each of the M subjects follows a common functional form, and the inter-individual differences arise due to population differences and different realizations of the Brownian motion paths $\{W_t^i\}_{t \geq t_0^i}$ and of the random parameters b^i.

5.1.3 Statistical Properties

- The solution to the quantities referenced in (5.15) is given by:

$$\begin{aligned}
\alpha(t,s) &= \theta(t-s), \\
-\beta(t,s) &= (\mu_0-\mu_d)(e^{\theta t}-e^{\theta s}) + \frac{\mu_d}{1-\frac{\delta}{\theta}}\left[e^{(\theta-\delta)t}-e^{(\theta-\delta)s}\right], \\
\gamma(t,s) &= \frac{\sigma^2}{4\theta}(e^{2\theta(t-s)}-1)
\end{aligned} \tag{5.19}$$

  Note that $s=0$ corresponds to the 48^{th} month challenge, and since subjects have their final challenge administered at randomized time-points, the above quantities evaluate to:

$$\begin{aligned}
\alpha(t^i,s=0) &= \theta t^i, \\
-\beta(t^i,s=0) &= (\mu_0-\mu_d)(e^{\theta t^i}-1) + \frac{\mu_d}{1-\frac{\delta}{\theta}}\left[e^{(1-\frac{\delta}{\theta})\theta t^i}-1\right], \\
\gamma(t^i,s=0) &= \frac{\sigma^2}{4\theta}(e^{2\theta t^i}-1)
\end{aligned} \tag{5.20}$$

- The conditional mean of the process is given by:

$$\begin{aligned}
E(X_t^i) &= e^{-\theta t^i}[\mu_0^i - \beta(t^i,0)] \\
&= \mu_0 - \mu_d(1-e^{Z^i\Gamma-\theta t^i}) + \frac{\mu_d}{1-\frac{\delta}{\theta}}\left[e^{-\delta t^i}-e^{-\theta t^i}\right] \\
&= \mu_0 - \mu_d(1-e^{Z^i\Gamma-\theta t^i}) + \frac{\mu_d e^{-\theta t^i}}{1-\frac{\delta}{\theta}}\left[e^{(1-\frac{\delta}{\theta})\theta t^i}-1\right]
\end{aligned} \tag{5.21}$$

  and the overall process mean across all individuals is given by:

$$E(X_{t^i}) = \mu_0 - \mu_d(1-e^{-\theta t^i}) + \frac{\mu_d e^{-\theta t^i}}{1-\frac{\delta}{\theta}}\left[e^{(1-\frac{\delta}{\theta})\theta t^i}-1\right] \tag{5.22}$$

- The process covariance is given by:

$$Cov(X_t^i, X_s^i) = e^{-\theta(s^i+t^i)}\left[\sigma_0^2 + \sigma_b^2 + \frac{\sigma^2}{2\theta}(e^{2\theta s}-1)\right] \tag{5.23}$$

Note that the Weiner process becomes stationary if $\sigma_0^2 + \sigma_b^2 = \frac{\sigma^2}{2\theta}$. In this situation,

$$\begin{aligned} Cov(X_t^i, X_s^i) &= \frac{\sigma^2}{2\theta} e^{-\theta(t^i - s^i)}, \\ Var(X_t^i) &= \frac{\sigma^2}{2\theta}, \\ \rho_{X_t^i, X_0^i} &= e^{-\theta(t^i - s^i)} \end{aligned} \tag{5.24}$$

where $\rho_{X_t^i, X_s^i}$ is the process correlation coefficient.

- Mean-Reversion Properties:

  Note that when $t^i \to \infty$, the overall process mean, $E(X_{t^i}) \to \mu_0 - \mu_d = \mu_\infty$ if $\frac{\delta}{\theta} > 1$ and θ is a finite positive constant. If $\frac{\delta}{\theta} \leq 1$, the process may not converge and will depend on the convergence rate of $\left[e^{-\delta t^i} - e^{-\theta t^i}\right]$.

5.2 Gauss-Markov (G-M) Process and Transition Density

5.2.1 Definition of G-M Process

A stochastic process $\{X_t, t \geq 0\}$ is called G-M process, if for a non-zero function h_t and non-decreasing function f_t,

1. $X_0 = 0$,
2. $\{X_t, t \geq 0\}$ has independent increments,
3. $X_t = h_t W(f_t)$, $\forall t > 0$, $W(*)$ is a standard Brownian motion derived from Wiener process, $W_t \sim N(0, t), \forall t > 0$.
4. The transition density for W_t is given by:

$$p(x_t, t | x_s, s) = \frac{1}{\sqrt{2\pi\sigma^2(t-s)}} \exp\big[- \frac{(x_t - x_s)^2}{2\sigma^2(t-s)}\big]$$

5.2.2 Properties of G-M Process

Few properties of the Gauss-Markov process of interest are:

- $E(X_t) = 0$ and $var(X_t) = h_t f_t$ where h_t, f_t could be chosen specific to application.

- X_t is a Markov process, allowing for memory-less property and exhibits fluctuating trajectories.
- G-M process is almost surely continuous without jumps. Therefore, it supposes no huge shocks in the system process.
- G-M process allow modeling nonlinear-variance, zero-mean Gaussian noises. Hence, it is natural to propose a drifted G-M X_t to model degradation process as follows:

$$X_t = q_t + h_t W(f_t),$$

with $E(X_t) = q_t$.

- If $f \in C\big[0,T\big]$, then the process defined by $Z_t = \int_0^t f_s dW_s$, $t \in \big[0,T\big]$ is a mean zero Gaussian process with independent increments and with covariance function:

$$Cov(Z_s, Z_t) = \int_0^{\min(s,t)} f_u^2 du.$$

- Note that Z_t can be treated as a Brownian motion under the time-change $Z_t = W(\rho_t)$ with $\rho_t := \int_0^t f_u^2 du$.

5.2.3 Connection between G-M process and Transition Density

Consider the general OU process X_t^i in (2.2.1) for an i^{th} subject, assuming initial threshold to be a constant X_s^i at time $s \geq 0$, the solution of which is expressed from (2.2.3) as:

$$X_t^i := X_{t,s(X_s^i)}^i = e^{-\alpha(t,s)}\big[X_s^i - \beta^i(t,s) + \int_s^t \sigma_u e^{\alpha(u,s)} dW_u^i\big]$$

then from (2.4.2) and (2.4.3), X_t^i is a drifted G-M process: $X_t^i = q_t^i + h_t W(f_t)$ under the time change $2\gamma(t,s) := \int_s^t \sigma_u^2 e^{2\alpha(u,s)} du$ where $q_{t,s} = e^{-\alpha(t,s)}\big[X_s^i - \beta^i(t,s)\big]$, $h_{t,s} = e^{-\alpha(t,s)}$ and $f_u = \sigma_u e^{\alpha(u,s)}$. This time-change technique can lead to the transition density for X_t^i. Note that the time-dependent OU process is G-M process only if the initial value X_0^i is deterministic.

5.3 Transition Density of OU based SDMEM

Since $f_u = \sigma_u e^{\alpha(u,s)}$ is deterministic and continuous, the stochastic integral $Z_{t,s} := \int_s^t \sigma_u e^{\alpha(u,s)} dW_u$, $t \geq s$ is a mean zero Gaussian process. Moreover, since $Z_{t,s} \sim N(0, 2\gamma(t,s))$, the transition density of $X^i_{t,s} = q^i_{t,s} + h_{t,s} Z_{t,s}$, given an initial threshold X^i_s at time $s \geq 0$, leads to:

$$p(x^i_t, t-s | x^i_s, b^i, \Gamma, \theta, \sigma) = N_{x^i_t | (x^i_s, b^i, \Gamma, \theta, \sigma)}(\mu^i_{ts}, \sigma^2_{ts}), \tag{5.25}$$

a Gaussian with mean $\mu^i_{ts} = e^{-\alpha(t,s)}\big(x^i_s - \beta^i(t,s)\big)$ and variance $\sigma^2_{ts} = 2\gamma(t,s)e^{-2\alpha(t,s)} = \frac{\sigma^2}{2\theta}\big[1 - e^{-2\theta(t-s)}\big]$, where $\alpha(t,s)$, $\beta^i(t,s)$, $\gamma(t,s)$ are as given in (5.19). For notational purposes, we will continue to ignore the fact that individuals are measured at different times until deemed necessary for computations.

5.3.1 Case when initial thresholds are random

In case of uncertain initial condition x^i_0 at $s = 0$ and assuming $x^i_0 \sim F_0(t) = N(\mu_0, \sigma_0)$, the conditional probability of transitioning to x^i_t at time t for i^{th} subject is given by:

$$\begin{aligned} p_{t0}(x^i_t, t | b^i, \Gamma, \theta, \sigma) &= \int_{-\infty}^{+\infty} p\left(x^i_t, t | u, b^i, \Gamma, \theta, \sigma\right) dF_0(u) \\ &= N(\mu_{t0}, \sigma_{t0}) \end{aligned} \tag{5.26}$$

where $\mu_{t0} = e^{-\theta t}\left[\mu_0 - \beta^i(t,0)\right]$ and $\sigma^2_{t0} = \left[\frac{\sigma^2}{2\theta}(1 - e^{-2\theta t})\right] + \sigma^2_0 e^{-2\theta t}$.

5.3.2 Case when Initial and Final Thresholds are Interval Censored

We note that subjects' true thresholds are interval censored due to the structure of food challenges with doses administered in $J = 6$ levels (Y = 100, 200, 500, 800, 1300, 2100). Let $[Y^i_{L_j}, Y^i_{L_{j+1}})$, $j = 1, ..., J$, represent an interval with $Y^i_{L_0} = 0, Y^i_{L_{J+1}} = \infty$ and correspondingly, $[X^i_{L_j}, X^i_{L_{j+1}})$ define the log transformed counterpart. Let $I^i_{tj} = I(X^i_{L_j} \leq X^i_t < X^i_{L_{j+1}})$ represent an indicator variable suggesting whether subject's true threshold x^i_t lies within the associated interval. Similarly, let $[X^i_{L_{j^*}}, X^i_{L_{j^*+1}})$ define the log transformed interval for initial threshold and $I^i_{0j^*} = I(X^i_{L_{j^*}} \leq X^i_0 < X^i_{L_{j^*+1}})$, the corresponding indicator. Then, the transition density of interval censored log thresholds at time t with interval censored random initial thresholds for i^{th} subject is extended as:

$$p_{ic}(x_t^i, t|b^i, \Gamma, \theta, \sigma) = \sum_{j^*=0}^{J}\sum_{j=0}^{J}\int_{X_{L_{j^*}}}^{X_{L_{j^*+1}}}\int_{X_{L_j}^i}^{X_{L_{j+1}}^i} I_{0j^*}^i I_{tj}^i N_{v|(u,s=0,b^i,\Gamma,\theta,\sigma)}(\mu_{t0}^i, \sigma_{t0}^2) N_u(\mu_0, \sigma_0^2) dv\, du$$
$$= \sum_{j^*=0}^{J}\sum_{j=0}^{J} I_{0j^*}^i I_{tj}^i \int_{X_{L_{j^*}}}^{X_{L_{j^*+1}}}\int_{X_{L_j}^i}^{X_{L_{j+1}}^i} ce^{-q^i(v,u)} dv\, du, \tag{5.27}$$

where the normalizing constant, $c = \left\{2\pi\sqrt{(1-\rho^2)}\sigma_v\sigma_u\right\}^{-1}$. Under the constraints of a stationary process, $\rho \equiv \rho_{X_t^i, X_0^i}|b_i = e^{-\theta t}$, $\sigma_v^2 = \sigma_u^2 = \frac{\sigma^2}{2\theta}$, from (5.20) - (5.24). Further, with the following defined quantities: $E(v) = e^{-\theta t}(\mu_0 - \beta^i(t,0))$ and $E(u) = \mu_0$, the exponent term $q^i(v,u)$ for a given b^i is a quadratic function of v and u given by:

$$q^i(v,u) = \frac{[v - E(v)]^2}{\sigma_v^2} + 2\rho\frac{\sigma_v}{\sigma_0}[v - E(v)](u - \mu_0) + \frac{(u-\mu_0)^2}{\sigma_0^2}$$

5.4 Maximum Likelihood

We assume that the distribution of $X_t^i \mid \left(X_s^i, X_0^i, b_i, \Theta = \{\Gamma, \delta, \theta, \sigma_b, \Omega\}, s < t\right)$, has a strictly positive density with respect to the Lebesgue measure on E and denote it by: $x \to p_X(x, t - s|x_s, x_0^i, \Theta) > 0, x \in E$. Here, Θ is a set of unknown parameters and hyper-parameters Ω. Generally, assume that subject i is observed $n_i + 1$ discrete time points $t_0^i, t_1^i, ..., t_{n_i}^i$, $i = 1, ..., M$. Let $\vec{x}^i$ be the vector of responses for subject i, $\vec{x}^i = x_0^i, x_1^i, ..., x_{n_i}^i$, where $x^i(t_j^i) = x_j^i$, and let $\vec{x} = (\vec{x}^1, ..., \vec{x}^M)$ be the N-dimensional total response vector, $N = \sum_i^M (n_i + 1)$. Define $\Delta_j^i = t_j^i - t_{j-1}^i$ for the time distance between the observations x_j^i and x_{j-1}^i. Then, the marginal density of $\vec{x}^i$ is obtained by integrating the conditional density of the data given the non-observable random effects b^i and Θ, with respect to their marginal densities, provided that W_t^i and b^i are independent of each other. This yields the likelihood function:

$$L(\Theta, b_i \mid X_0^i, X_s^i) = \prod_i^M \int_B \int_\Theta p_{\vec{X}}(\vec{x}^i|b^i, \Theta)\, p_B(b^i|\sigma_b)\, p_\Theta\, db^i d\Theta \tag{5.28}$$

where the density functions:

$$
\begin{aligned}
p_B(b^i|\sigma_b) &= N(0,\sigma_b) \\
p_{\vec{X}}(\vec{x}^i|b^i,\Theta) &= \prod_{j=1}^{n_i} p_X(x_j^i,\Delta_j^i|x_{j-1}^i,b^i,\Theta) \\
&= \{\prod_{j=2}^{n_i} p_X(x_j^i,\Delta_j^i|x_{j-1}^i,b^i,\Theta)\} * p_{10}(x_1^i,\Delta_j^i|b^i,\Theta)
\end{aligned}
$$

Note that $p_X(x_j^i,\Delta_j^i|x_{j-1}^i,b^i,\Theta)$ is a transition density given by (5.25) with $t=t_j^i$ and $s=t_{j-1}^i$. And, $p_{10}(x_1^i,\Delta_j^i|b^i,\Theta)$ is a transition density given by (5.26) with $t=t_1^i$.

However, in case of our application, each subject is observed only twice: at the 48^{th} month DBPCFC, $t_0=0$ and at a randomized time t_i of final DBPCFC. Further, the true thresholds at at these time-points are interval censored due to the structure of food challenges, Under this scenario, we directly use (5.27) towards computing the likelihood. This yields the following likelihood function:

$$
\begin{aligned}
L(\Theta|X_{L_j}^i,X_{L_{j+1}}^i,X_{L_{j'}}^i,X_{L_{j'+1}}^i) &= \prod_i^M \int_B \int_\Theta p_{ic}(x_t^i,t|b^i,\Gamma,\theta,\sigma)\, p_B(b^i|\tau)\, p_\Theta\, db^i\, d\Theta \\
&= \int_\Theta \frac{c}{2\pi\sqrt{\sigma_b^2}} \prod_i^M \sum_{j^*=0}^J \sum_{j=0}^J I_{0j^*}^i I_{tj}^i \int_{X_{L_{j^*}}^i}^{X_{L_{j^*+1}}^i} \int_{X_{L_j}^i}^{X_{L_{j+1}}^i} \int_B \\
&\qquad \exp\left(-q^i(v,u) - \frac{{b^i}^2}{2\sigma_b^2}\right) db^i\, dv\, du\, p_\Theta\, d\Theta
\end{aligned}
\tag{5.29}
$$

Generally, there is no closed form for the likelihood function $L(\Theta)$. Due to the analytic intractability of SDEs regulating most nonlinear multivariate diffusions, the likelihood-based inference methods can be problematic. To overcome this challenge, few methods have been proposed that include closed-form expansion of the transition density [1; 2; 75], exact simulation approaches [6; 70] and use of the Euler-Maruyama approximation coupled with data augmentation [24; 29; 74; 43]. There has been relatively little work on SDMEMs due to the challenges that inference for SDEs present. On the other side, [21] discuss inference for SDMEMs in a Bayesian framework, and implement a Gibbs sampler when the SDE (for each experimental unit) has an explicit solution. When no explicit solution exists they suggest that a solution might be found using the Euler-Maruyama discretization.

5.5 Bayesian Analysis

The Bayesian MCMC method that obtains posterior distributions of the parameters is a useful alternative to perform inference on SDEs. The procedure's validity rests on a careful analysis of the stationary distribution of the chain, which involves an extension of the theory of [56]. Further, [46] contend that the "posterior" distribution of the profile likelihood with respect to a prior on Θ is asymptotically equivalent to the distribution of the maximum profile likelihood estimator $\hat{\Theta}$ and that inferences about Θ might also be based on the marginal posterior of Θ from the full likelihood with respect to a joint prior on (Θ, η) where η is a nuisance parameter. Thus, the inference of the parameters for the Bayesian method can be based on their posterior distributions.

For the class of SDMEM considered in this chapter, that uses incomplete discrete-time observations made sparsely in time and is subject to censoring (so that only a subset of model components are observed), we resort to Bayesian inference borrowing ideas from [88; 90]. Although a discretization bias is introduced, this is made arbitrarily small (at greater computational expense). A Bayesian approach then aims to construct the joint posterior density for parameters and the components of the latent process. The intractability of the posterior density necessitates simulation techniques such as Markov chain Monte Carlo. As is well documented in [65], care must be taken in the design of the MCMC sampler due to dependence between the parameters entering the population mean and the latent process. To capture nonlinear dynamics exhibited between observation times, a key requirement is the ability to sample the latent process between two fixed values. For this, we borrow strength of the study design that allows latter value to vary across individuals, thus enabling information across the entire time spectrum of the study.

Essentially, a hybrid Metropolis-Hastings scheme is implemented for the interval-censored threshold data to calculate the marginal likelihood of all parameters of interest, targeting their posterior distribution. It involves generating a Markov chain $\{\Theta(1), \Theta(2), ...\}$ with stationary density proportional to $p_{(\Theta,n)}(\Theta) = l_n(\Theta)q(\Theta)$, where $q(\Theta) = Q(d\Theta)/(d\Theta)$ for some prior measure Q. The procedure begins with an initial value $\Theta(1)$ for the chain. For each $k = 2, 3, ...$, a proposal $\vartheta(k+1)$ is obtained by random walk from $\Theta(k)$. The quantities, $\hat{\eta}_{\vartheta^{(k+1)}}$ and $p_{\vartheta^{(k+1)},n}(\vartheta^{(k+1)})$, are computed and based on acceptance rule, a decision is made to accept $\vartheta^{(k+1)}$ by evaluating the ratio $p_{\vartheta^{(k+1)},n}(\vartheta^{(k+1)})/p_{\vartheta^{(k)},n}(\vartheta^{(k)})$. After generating a sufficiently long chain, the mean of the chain is

computed to estimate the maximizer of $l_n(\Theta)$ and the variance to estimate the inverse Information matrix.

This algorithm is implemented in the publicly available software WinBUGS [73]. While building the model, a bivariate Gaussian distribution for the initial and latter interval-censored log thresholds is specified as in (5.27). Censoring is specified such that individual's latent log thresholds are bound by the observed lower and upper limit to contribute to the full conditional distribution. Priors are then assigned to the unknown independent parameters. Regression coefficients, γ_1 and γ_2, are assigned uninformative standard normal priors with $zero$ mean and the intercept, γ_0 is constrained to allow $E(\exp Z^i\Gamma) = 1$. Uniform prior is assumed for the drift parameter $\theta \sim U(0, 0.5)$ and the population drift parameter is re-parametrized in terms of the ratio of population and process drift parameters, $\frac{\delta}{\theta} \sim U(0, 10)$, which is constrained to not equal 1. The priors for the variance parameters, σ_0^2 and σ_b^2, are assumed to be *Inverse Gamma(shape = a, rate = b)* with hyper-parameters (a, b) chosen such that the priors are weakly informative. For example, for σ_0^2, we set $a = b = 0.0001$ and for σ_b^2, we set $a = 5$, $b = 0.01$. Due to stationarity of the latent process, we induce inverse relationship between σ_0^2 and σ_b^2. For this reason, we allow uninformative prior on σ_0^2 to cover a spectrum of challenge thresholds than those observed, thus allowing it to reflect variability in observed covariates. Further, the inherent individual's variability, σ_b^2, that remains unexplained by observed covariates is constrained to be much smaller. This approach relaxes the inverse relationship and makes estimation of the two parameters independent of each other. Parameter estimates and the 95% credible intervals (CI) can be obtained from the MCMC samples.

Some caution is exercised because the choice of hyper-parameters may affect the final parameter estimates. Initial values for the parameters are selected within ± 2 standard deviations of the prior means. To assess the stability of final estimates in our simulations, two parallel chains are run with different initial values. The Gaussian proposal distribution used for this algorithm adapts for the first 4000 iterations and these samples are discarded from the summary statistics. A further 1000 update burn-in followed by 15000 updates were considered to give final parameter estimates. The convergence of the MCMC samples of optimization parameters is visually inspected and confirmed using the criteria of [10]. Sample WinBUGS code is shown in the appendix.

5.6 Simulations: Finite Sample Performance

For simulating the data, the true conditional distribution of initial and final challenge log thresholds, $\mathbf{X^i} = (X_0^i, X_2^i)'$ is specified to be $N_2(\mu^{\mathbf{i}}, \mathbf{\Sigma^i_{0,t}})$ with conditional mean vector, $\mu^{\mathbf{i}} = \left(\mu_0^i, E(X_t^i)\right)'$ and covariance structure,

$$\mathbf{\Sigma^i_{0,t}} = \frac{\sigma^2}{2\theta}\begin{pmatrix} 1 & e^{-\theta t^i} \\ e^{-\theta t^i} & 1 \end{pmatrix}$$

where $\mu_0^i = (\mu_0 - \mu_d) + \mu_d e^{Z^i\Gamma}$, $E(X_t^i)$ is as given by (5.21) and $\sigma^2 = 2\theta(\sigma_0^2 + \sigma_b^2)$ to ensure that the process is stationary. Subjects are allocated their randomized times, t^i as per the study protocol described in chapter 4. Baseline covariates, $\mathbf{Z} = (Z_1, Z_2)'$ are simulated such that $Z_1 \sim \text{N}(0,1)$, $Z_2 \sim$ Bernoulli$(p = 0.7)$ and three settings of corresponding regression coefficients, $\mathbf{\Gamma} = (\gamma_1, \gamma_2)' \in \{(0.05, -0.05)', (0.1, -0.1)', (0.5, -0.5)'\}$ are considered. The intercept coefficient, γ_0 is restricted to equal $-\left(\gamma_1^2/2 + \log(pe^{\gamma_2} + 1 - p)\right)$ to ensure $E(e^{Z^i\Gamma}) = 1$ and σ_0^2 is empirically evaluated to equal $\left(\mu_d^2 Var(e^{Z'\Gamma})\right)$ to reflect variability in covariates. Here, we focus on the Γ parameters that represent the effect of covariates on expected log-threshold at any time-point t where a positive (negative) coefficient is associated with an increase (decrease) in log threshold. Further, $\mu_0 = 7.8$, $\mu_d = 3.4$, $\theta = 0.01$, $\frac{\delta}{\theta} = 5$ and $\sigma_b = 0.05$ are chosen a priori to reflect potential data arising from the trial. Finally, the challenge threshold intervals for a subject are evaluated based on the pre-decided challenge dose levels the simulated thresholds originating from the latent process fall in. Due to high variability attributed to larger regression coefficients, it is likely that the thresholds at time t^i simulated for the latter challenge fall in interval higher than that at the initial challenge. In such cases, the latter threshold intervals are amended to $(LOAEL^i_{t=0}, \infty)$, since subjects in the actual trial are not administered challenge doses above their initial LOAEL levels and there is no way to assess where their potential threshold would lie in the spectrum. We then perform the following two numerical experiments to investigate the finite sample performance of the proposed procedure within context of the data expected from the study.

5.6.1 Experiment-I: Varying Effect Sizes

In the first setting, the above configuration is adopted for two sample sizes, $M = 51$, the actual study sample size, and $M = 200$ as a benchmark for method's performance under large sample size. This procedure is repeated for another dataset generated using a different seed, leading to twelve simulations in total, six under each seed. The change in seed allow the set of covariates to differ for a given set of parameters. This elucidates procedure's consistency under two different set of covariates for small and large sample sizes.

Table 5.1 summarizes parameter estimates and corresponding 95% Bayesian Credible Intervals (BCI). In general, we observe more precise estimates with narrower intervals for higher sample size unless more censoring is involved. As depicted by Scenario-1 where there is no censoring, the higher sample size gives parameter estimates with little bias; on the other hand, the estimates from lower sample size exhibit some bias but are in the right direction in vicinity of the truth with wider intervals as expected. The estimates of $(\gamma_1, \gamma_2, \mu_d)$ appear more sensitive to differences in data. As evident by Scenario-2, higher censoring rate is associated with larger bias when comparing datasets with the same sample size but different seeds. Finally, larger covariate coefficients depicting absolute effect aggravate bias due to higher censoring rate, as shown in Scenario 3. A large coefficient induces more variability in threshold distribution leading to heavier tails beyond observation limits.

	Seed 1				Seed 2			
	M = 51		M = 200		M = 51		M = 200	
Node	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**
SCENARIO 1								
$\gamma_1(0.05)$	0.08 (0.015)	(0.05, 0.11)	0.05 (0.007)	(0.04, 0.07)	0.06 (0.014)	(0.03, 0.09)	0.05 (0.007)	(0.04, 0.06)
$\gamma_2(-0.05)$	-0.08 (0.029)	(-0.13, -0.02)	-0.05 (0.012)	(-0.07, -0.02)	-0.03 (0.023)	(-0.08, 0.01)	-0.04 (0.011)	(-0.06, -0.02)
$\delta(0.05)$	0.06 (0.012)	(0.04, 0.09)	0.05 (0.006)	(0.04, 0.06)	0.05 (0.01)	(0.03, 0.07)	0.05 (0.006)	(0.04, 0.06)
$\theta(0.01)$	0.01 (0.002)	(0.01, 0.02)	0.01 (0.001)	(0.01, 0.01)	0.01 (0.002)	(0.01, 0.02)	0.01 (0.001)	(0.01, 0.01)
$\mu_0(7.8)$	7.81 (0.037)	(7.74, 7.88)	7.8 (0.017)	(7.77, 7.84)	7.77 (0.038)	(7.7, 7.84)	7.78 (0.018)	(7.74, 7.82)
$\mu_d(3.4)$	2.79 (0.295)	(2.27, 3.43)	3.26 (0.241)	(2.84, 3.79)	3.48 (0.553)	(2.58, 4.82)	3.59 (0.345)	(2.99, 4.26)
$\sigma_0(0.19)$	0.19 (0.025)	(0.15, 0.25)	0.18 (0.012)	(0.15, 0.2)	0.2 (0.027)	(0.16, 0.26)	0.2 (0.014)	(0.17, 0.22)
$\sigma_b(0.05)$	0.05 (0.012)	(0.03, 0.08)	0.05 (0.011)	(0.03, 0.07)	0.05 (0.012)	(0.03, 0.08)	0.05 (0.014)	(0.03, 0.08)
Censoring (N) (L_0, R_0, L_t, R_t)	(0, 0, 0, 0)		(0, 0, 0, 0)		(0, 0, 0, 0)		(0, 0, 0, 0)	
Threshold ↑ (N)	0		1		0		0	
SCENARIO 2								
$\gamma_1(0.1)$	0.11 (0.031)	(0.06, 0.18)	0.11 (0.014)	(0.08, 0.14)	0.1 (0.024)	(0.06, 0.15)	0.08 (0.012)	(0.06, 0.1)
$\gamma_2(-0.1)$	-0.06 (0.046)	(-0.16, 0.02)	-0.07 (0.02)	(-0.11, -0.03)	-0.09 (0.037)	(-0.18, -0.03)	-0.09 (0.019)	(-0.13, -0.05)
$\delta(0.05)$	0.05 (0.013)	(0.03, 0.08)	0.05 (0.008)	(0.04, 0.07)	0.05 (0.012)	(0.03, 0.07)	0.05 (0.007)	(0.03, 0.06)
$\theta(0.01)$	0.01 (0.002)	(0.01, 0.02)	0.01 (0.001)	(0.01, 0.01)	0.01 (0.002)	(0.01, 0.01)	0.01 (0.001)	(0.01, 0.01)
$\mu_0(7.79)$	7.79 (0.067)	(7.66, 7.92)	7.77 (0.031)	(7.71, 7.83)	7.69 (0.057)	(7.58, 7.8)	7.77 (0.031)	(7.71, 7.83)
$\mu_d(3.4)$	3.37 (0.708)	(2.28, 4.97)	3.54 (0.336)	(2.96, 4.27)	3.66 (0.736)	(2.5, 5.46)	3.88 (0.468)	(3.09, 4.95)
$\sigma_0(0.38)$	0.44 (0.041)	(0.36, 0.52)	0.39 (0.019)	(0.36, 0.43)	0.36 (0.038)	(0.29, 0.44)	0.39 (0.02)	(0.36, 0.43)
$\sigma_b(0.05)$	0.05 (0.013)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)	0.05 (0.013)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)
Censoring (N) (L_0, R_0, L_t, R_t)	(0, 7, 0, 1)		(0, 23, 0, 3)		(0, 4, 0, 0)		(0, 14, 0, 0)	
Threshold ↑ (N)	1		2		0		2	
SCENARIO 3								
$\gamma_1(0.5)$	0.43 (0.158)	(0.17, 0.78)	0.32 (0.06)	(0.22, 0.45)	0.37 (0.146)	(0.14, 0.71)	0.33 (0.069)	(0.21, 0.48)
$\gamma_1(-0.5)$	-0.36 (0.207)	(-0.83, -0.02)	-0.22 (0.079)	(-0.39, -0.08)	-0.31 (0.186)	(-0.74, -0.01)	-0.35 (0.091)	(-0.55, -0.19)
$\delta(0.05)$	0.04 (0.023)	(0.01, 0.09)	0.03 (0.009)	(0.02, 0.05)	0.05 (0.026)	(0.01, 0.1)	0.04 (0.01)	(0.02, 0.06)
$\theta(0.01)$	0.01 (0.002)	(0, 0.01)	0.01 (0.001)	(0.01, 0.01)	0.01 (0.003)	(0, 0.01)	0.01 (0.001)	(0.01, 0.01)
$\mu_0(7.74/7.5)$	7.28 (0.299)	(6.69, 7.86)	7.39 (0.146)	(7.11, 7.68)	7.25 (0.306)	(6.64, 7.85)	7.58 (0.157)	(7.28, 7.89)
$\mu_d(3.4)$	4.17 (2.058)	(1.89, 9.13)	4.73 (0.905)	(3.25, 6.75)	4.63 (2.17)	(1.96, 10.21)	4.78 (0.911)	(3.31, 6.81)
$\sigma_0(2.05)$	1.95 (0.216)	(1.58, 2.42)	1.89 (0.101)	(1.7, 2.09)	2.01 (0.206)	(1.65, 2.45)	2.02 (0.106)	(1.82, 2.24)
$\sigma_b(0.05)$	0.05 (0.012)	(0.03, 0.08)	0.05 (0.013)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)	0.05 (0.013)	(0.03, 0.08)
Censoring (N) (L_0, R_0, L_t, R_t)	(10, 19, 8, 12)		(31, 64, 42, 43)		(10, 15, 13, 7)		(30, 70, 50, 37)	
Threshold ↑ (N)	13		40		10		36	

Table 5.1: Experiment-I: Summary includes the mean parameter estimates and 95% Bayesian Credible Intervals (BCI) across all considered configurations.

5.6.2 Experiment-II: Fixed Effect Size, Different Datasets with Varying Covariates

In the second setting, we fix $\mathbf{\Gamma} = (\gamma_1, \gamma_2)' = (0.1, -0.1)'$ and for each of the two sample sizes, $M = \{51, 200\}$, we perform three additional simulations using different seeds to total five simulations corresponding to data sets with differing covariates originating from the same set of parameters. This ensures consistency of the procedure under variety of covariate data with regards to finite sample performance.

Based on prior information and the data collected thus far, this choice of $\mathbf{\Gamma}$ is expected to represent study data more closely with regards to the level of left/right censoring between $2 - 18\%$ on both the challenge tests. While unlikely in reality, it also presented us a case when final challenge thresholds could be higher than the initial. Table 5.2 provides a summary of the estimated parameters that had priors assigned. The summary includes posterior mean, estimated standard errors and 95% BCI. Convergence is achieved for all parameters and their estimates are not sensitive to the choice of hyper-parameters and initial values. From these results, one will note that the bias is generally small and the BCI include the corresponding true parameters for all the datasets. The estimated parameters for the two sample sizes within the same seed are quite close with wider BCI for smaller sample size. In general, the posterior regression coefficient estimates are closest to the truth when the overall left/right censoring rate is the lowest, as depicted by Seed 2.

Further, we assess the empirical goodness of model fit by comparing temporal curves of the estimated mean log-thresholds as described by the latent process with that assumed. The curves, as shown in Figure 5.1 are predictions based on all subjects without adjusting for covariates. Each curve corresponds to the posterior distribution of expected log-thresholds with solid red depicting median and dashed blue, the 95% posterior interval. The curve plotted in black dots corresponds to the assumed parametric values and represents the truth. We observe that in majority of the cases, the small sample size does reasonably well in predicting expected log-threshold. When compared to the larger sample size, they have wider CI at later time points where we lack data. This observation is not unexpected.

In summary, the two numerical experiments illustrate finite sample performance of the proposed methodology in context of an innovative study design that is embarked on a typical interval censored data from food challenges administered at randomized time-points. We have shown that this

methodology could be used to study covariate effects on expected threshold trajectory, enabling inferences on expected tolerance threshold at a specific time-point or time to reach an arbitrary threshold with some added work. The procedure does reasonably well with small sample size when left/right censoring is moderate.

	M = 51									
	Seed 1		Seed 2		Seed 3		Seed 4		Seed 5	
Node	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**	**Mean (SD)**	**95% BCI**
$\gamma_1(\mathbf{0.10})$	0.11 (0.031)	(0.06, 0.18)	0.09 (0.023)	(0.05, 0.14)	0.16 (0.042)	(0.08, 0.24)	0.08 (0.022)	(0.04, 0.13)	0.19 (0.043)	(0.11, 0.28)
$\gamma_2(-\mathbf{0.10})$	-0.06 (0.046)	(-0.16, 0.02)	-0.09 (0.035)	(-0.17, -0.03)	-0.16 (0.077)	(-0.32, -0.02)	-0.12 (0.039)	(-0.21, -0.05)	-0.12 (0.055)	(-0.24, -0.02)
$\delta(\mathbf{0.05})$	0.05 (0.013)	(0.03, 0.08)	0.06 (0.016)	(0.03, 0.09)	0.07 (0.019)	(0.04, 0.12)	0.06 (0.015)	(0.04, 0.1)	0.06 (0.016)	(0.04, 0.1)
$\theta(\mathbf{0.01})$	0.01 (0.002)	(0.01, 0.02)	0.01 (0.002)	(0, 0.01)	0.01 (0.003)	(0.01, 0.02)	0.01 (0.003)	(0.01, 0.02)	0.01 (0.003)	(0.01, 0.02)
$\mu_0(\mathbf{7.79})$	7.79 (0.067)	(7.66, 7.92)	7.69 (0.057)	(7.58, 7.8)	7.93 (0.062)	(7.81, 8.05)	7.86 (0.05)	(7.76, 7.96)	7.9 (0.066)	(7.78, 8.03)
$\mu_\mathbf{d}(\mathbf{3.40})$	3.37 (0.708)	(2.28, 4.97)	3.85 (0.822)	(2.56, 5.84)	2.53 (0.419)	(1.87, 3.53)	2.99 (0.5)	(2.21, 4.19)	3.02 (0.531)	(2.22, 4.28)
$\sigma_0(\mathbf{0.38})$	0.44 (0.041)	(0.36, 0.52)	0.36 (0.038)	(0.29, 0.44)	0.39 (0.038)	(0.32, 0.47)	0.31 (0.033)	(0.25, 0.38)	0.41 (0.041)	(0.34, 0.5)
$\sigma_\mathbf{b}(\mathbf{0.05})$	0.05 (0.013)	(0.03, 0.08)	0.05 (0.013)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)	0.05 (0.013)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)
Censoring (N)										
$(\mathbf{L_0}, \mathbf{R_0}, \mathbf{L_t}, \mathbf{R_t})$	(0, 7, 0, 1)		(0, 4, 0, 0)		(0, 8, 0, 1)		(0, 6, 0, 0)		(0, 9, 1, 3)	
Threshold ↑ (N)	1		0		1		0		0	

	M = 200									
	Simulation 1		Simulation 2		Simulation 3		Simulation 4		Simulation 5	
Node	**Mean (SD)**	**95% CI**	**Mean (SD)**	**95% CI**	**Mean (SD)**	**95% CI**	**Mean (SD)**	**95% CI**	**Mean (SD)**	**95% CI**
$\gamma_1(\mathbf{0.10})$	0.11 (0.014)	(0.08, 0.14)	0.08 (0.012)	(0.06, 0.1)	0.09 (0.012)	(0.07, 0.12)	0.09 (0.014)	(0.06, 0.11)	0.1 (0.014)	(0.08, 0.13)
$\gamma_2(-\mathbf{0.10})$	-0.07 (0.02)	(-0.11, -0.03)	-0.09 (0.019)	(-0.13, -0.05)	-0.1 (0.02)	(-0.14, -0.06)	-0.09 (0.02)	(-0.14, -0.06)	-0.09 (0.021)	(-0.14, -0.06)
$\delta(\mathbf{0.05})$	0.05 (0.008)	(0.04, 0.07)	0.05 (0.007)	(0.03, 0.06)	0.07 (0.013)	(0.05, 0.1)	0.06 (0.011)	(0.04, 0.08)	0.06 (0.012)	(0.04, 0.08)
$\theta(\mathbf{0.01})$	0.01 (0.001)	(0.01, 0.01)	0.01 (0.001)	(0.01, 0.01)	0.01 (0.001)	(0.01, 0.01)	0.01 (0.002)	(0.01, 0.01)	0.01 (0.001)	(0.01, 0.01)
$\mu_0(\mathbf{7.79})$	7.77 (0.031)	(7.71, 7.83)	7.77 (0.031)	(7.71, 7.83)	7.87 (0.028)	(7.82, 7.93)	7.84 (0.029)	(7.78, 7.9)	7.81 (0.031)	(7.75, 7.87)
$\mu_\mathbf{d}(\mathbf{3.40})$	3.54 (0.336)	(2.96, 4.27)	3.88 (0.468)	(3.09, 4.95)	3.39 (0.317)	(2.86, 4.1)	3.69 (0.42)	(2.95, 4.56)	3.81 (0.389)	(3.15, 4.67)
$\sigma_0(\mathbf{0.38})$	0.39 (0.019)	(0.36, 0.43)	0.39 (0.02)	(0.36, 0.43)	0.36 (0.018)	(0.32, 0.4)	0.37 (0.019)	(0.34, 0.41)	0.4 (0.02)	(0.36, 0.44)
$\sigma_\mathbf{b}(\mathbf{0.05})$	0.05 (0.012)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)	0.05 (0.013)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)	0.05 (0.012)	(0.03, 0.08)
Censoring (N)										
$(\mathbf{L_0}, \mathbf{R_0}, \mathbf{L_t}, \mathbf{R_t})$	(0, 23, 0, 3)		(0, 14, 0, 0)		(0, 21, 0, 2)		(0, 19, 0, 6)		(0, 21, 1, 5)	
Threshold ↑ (N)	2		2		2		2		1	

Table 5.2: Experiment-II: Summary includes the mean parameter estimates and 95% Bayesian Credible Intervals (BCI) using data generated from five different seeds with the same set of parameters and $\Gamma = (0.1, -0.1)'$.

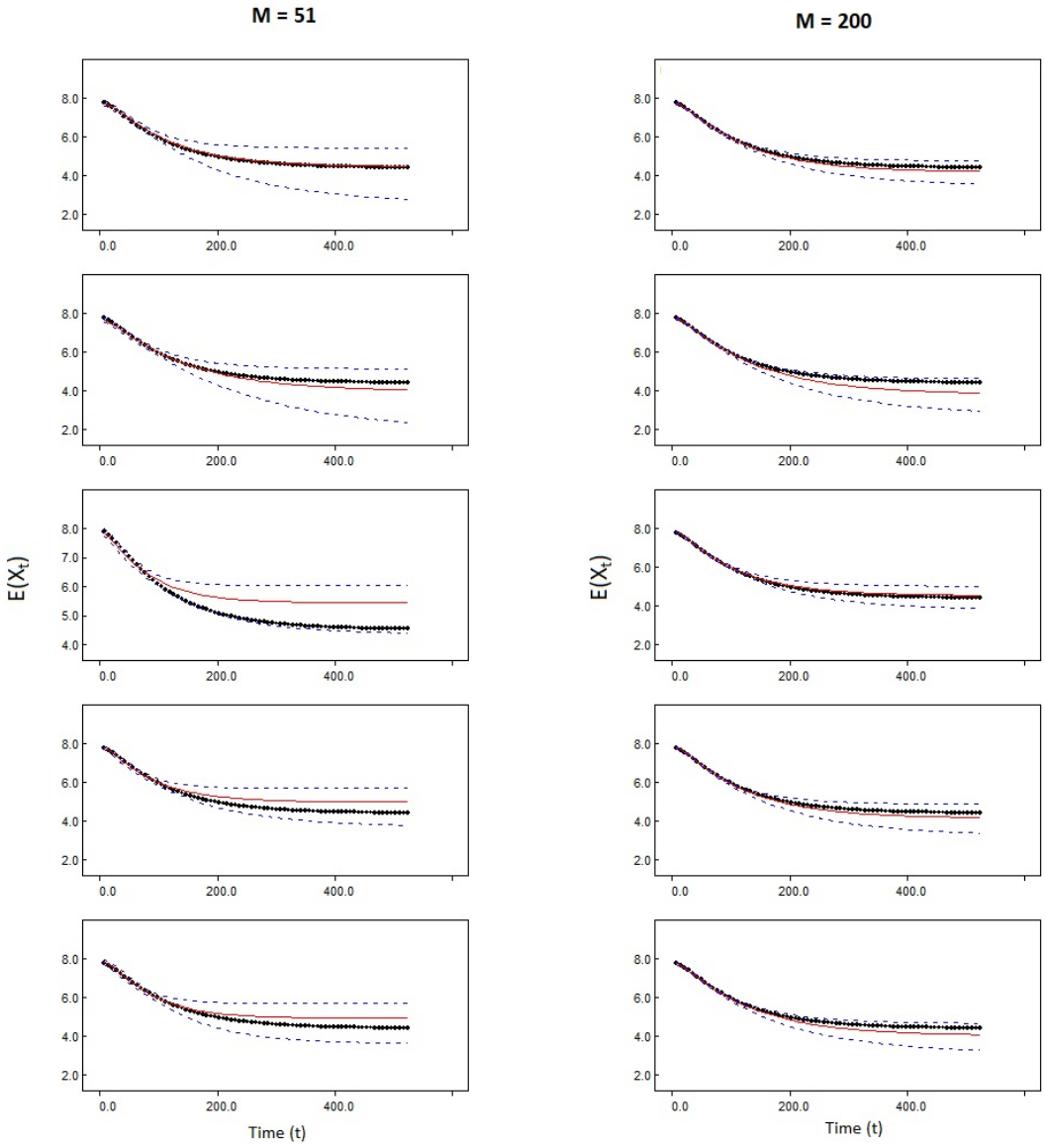

Figure 5.1: This figure provides log-threshold projections for the two sample sizes, $M = \{51, 200\}$, derived from Experiment-II. Each curve corresponds to the posterior distribution of expected log-thresholds with solid red depicting median and dashed blue, the 95% posterior interval. The curve plotted in black dots corresponds to the assumed parametric values and represents the truth.

5.7 Discussion

In this chapter, we focused on developing and validating a flexible methodology guided by a scientifically motivated stochastic process under the mixed effects framework. The method builds on the bulwark of the study design described in Chapter 2 by delineating individual trajectories from data collected at random time-points and thus maximizing information from an otherwise typically sparse food challenge study that administers challenges at two fixed occasions. To this end, this work has proposed a model, under the assumptions of a latent process, which can be used to analyze covariate effects on threshold projections. Through the model development, we adopt Bayesian approach for estimating the parameters underlying the model. Primarily, a hybrid Metropolis-Hastings algorithm is implemented within a Gibbs sampler until convergence is achieved on all the unknown parameters. The resulting algorithm is flexible and easy to implement and the estimates are independent of the choice of hyper-parameters and initial values.

The finite sample performance of the proposed approach was exhibited through extensive numerical study. This study suggests that under moderate rate of censoring on extreme ends of threshold distribution, the procedure performs fairly well in estimating parameters with low sample size. Since the food challenge studies are typically small, a study can benefit from a well-designed challenge built on a priori scientific information to form threshold intervals. Overall, this methodological approach of assessing covariate effects on time-bound threshold decay is promising and could be further extended or adapted.

5.8 Future Directions

In parlance of systematic immunology, the concept of FPT borrowed from reliability studies is the time taken for the immune system to fail when the tolerance threshold deteriorates beyond a preset level. Mathematically, if a stochastic degradation process Y_t has been established, it would be convenient to consider the failure time as the random time $\tau = inf_{t>0}\{Y_t \geq D(t)\}$ where $D(t)$ is some specified threshold of interest [18]. Once the process parameters are estimated, one can apply direct maximization techniques, e.g., the Newton-Raphson method or the Expectation-Maximization (EM) algorithm, to estimate FPT. In a Bayesian framework, based on the posterior distribution of expected log-thresholds, one can empirically measure the earliest time an estimated trajectory hits a

specified threshold and derive posterior distribution of FPT which could then be used for inferential purposes. Since it is possible to estimate an individual or population level curve, one can draw useful inferences on "$half - life$" of population sensitivity threshold or individual's expected threshold dropping by at least a level of the administered dose or even below MCRT.

Using a more flexible semi-parametric approach, the methodology could further be extended to allow individual specific drift and diffusion parameters be functions of covariates and vary with time. Similar estimation approach could be used to derive inferences based on the profile likelihood at an individual as well as population level. To reduce computation burden, the model parameters and latent trajectories could be updated jointly [30] and their conditional dependencies could be dealt using blocking strategies as proposed by [88].

The flexibility that these models offer could be put to clinical use to inform individualized treatments. Based on the study of risk factors and prognosis of individual SU failure, preventive strategies could be adopted in practice and more targeted treatment could be administered.

BIBLIOGRAPHY

[1] Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica*, 70(1):223–262.

[2] Ait-Sahalia, Y. et al. (2008). Closed-form likelihood expansions for multivariate diffusions. *The Annals of Statistics*, 36(2):906–937.

[3] Akinbo, B., Faniran, T., and Ayoola, E. (2015). Numerical solution of stochastic differential equations. *International Journal of Advanced Research in Science, Engineering and Technology*, 2:608–616.

[4] Allison, P. D. (2010). *Survival analysis using SAS: a practical guide*. Sas Institute.

[5] Baumert, J. L., Martin, L., Thébault, C., Taylor, S. L., Koppelman, S. J., and Ruban, C. (2016). Quantitative assessment of the safety benefits associated with increasing clinical peanut thresholds through immunotherapy. *Journal of Allergy and Clinical Immunology*, 137(2):AB126.

[6] Beskos, A., Papaspiliopoulos, O., Roberts, G. O., and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(3):333–382.

[7] Betensky, R. A., Lindsey, J. C., Ryan, L. M., and Wand, M. (2002). A local likelihood proportional hazards model for interval censored data. *Statistics in Medicine*, 21(2):263–275.

[8] Bock, S. A., Muñoz-Furlong, A., and Sampson, H. A. (2007). Further fatalities caused by anaphylactic reactions to food, 2001-2006. *Journal of Allergy and Clinical Immunology*, 119(4):1016–1018.

[9] Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American statistical Association*, 88(421):9–25.

[10] Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*, 7(4):434–455.

[11] Burks, A. W., Jones, S. M., Wood, R. A., Fleischer, D. M., Sicherer, S. H., Lindblad, R. W., Stablein, D., Henning, A. K., Vickery, B. P., Liu, A. H., et al. (2012). Oral immunotherapy for treatment of egg allergy in children. *New England Journal of Medicine*, 367(3):233–243.

[12] Burks, A. W., Wood, R. A., Jones, S. M., Sicherer, S. H., Fleischer, D. M., Scurlock, A. M., Vickery, B. P., Liu, A. H., Henning, A. K., Lindblad, R., et al. (2015). Sublingual immunotherapy for peanut allergy: long-term follow-up of a randomized multicenter trial. *Journal of Allergy and Clinical Immunology*, 135(5):1240–1248.

[13] Cai, B., Lin, X., and Wang, L. (2011). Bayesian proportional hazards model for current status data with monotone splines. *Computational Statistics & Data Analysis*, 55(9):2644–2651.

[14] Cai, T. and Betensky, R. A. (2003). Hazard regression for interval-censored data with penalized spline. *Biometrics*, 59(3):570–579.

[15] Carey, V. and Wang, Y.-G. (2001). Mixed-effects models in s and s-plus.

[16] Chen, C.-M., Lai, C.-C., Cheng, K.-C., Weng, S.-F., Liu, W.-L., and Shen, H.-N. (2015). Effect of end-stage renal disease on long-term survival after a first-ever mechanical ventilation: a population-based study. *Critical Care*, 19(1):354.

[17] Cox, D. R. (1992). Regression models and life-tables. In *Breakthroughs in statistics*, pages 527–541. Springer.

[18] Deng, Y., Barros, A., and Grall, A. (2016). Degradation modeling based on a time-dependent ornstein-uhlenbeck process and residual useful lifetime estimation. *IEEE Transactions on Reliability*, 65(1):126–140.

[19] Ditlevsen, S. and De Gaetano, A. (2005). Mixed effects in stochastic differential equation models. *REVSTAT-Statistical Journal*, 3(2):137–153.

[20] Ditlevsen, S. and Sørensen, M. (2004). Inference for observations of integrated diffusion processes. *Scandinavian Journal of Statistics*, 31(3):417–429.

[21] Donnet, S., Foulley, J.-L., and Samson, A. (2010). Bayesian analysis of growth curves using mixed models defined by stochastic differential equations. *Biometrics*, 66(3):733–741.

[22] Donnet, S. and Samson, A. (2008). Parametric inference for mixed models defined by stochastic differential equations. *ESAIM: Probability and Statistics*, 12:196–218.

[23] Dorey, F. J., Little, R. J., and Schenker, N. (1993). Multiple imputation for threshold-crossing data with interval censoring. *Statistics in medicine*, 12(17):1589–1603.

[24] Durham, G. B. and Gallant, A. R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business & Economic Statistics*, 20(3):297–338.

[25] Finkelstein, D. M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, pages 845–854.

[26] for Complementary, N. C. and of North Carolina at Chapel Hill, I. H. N. U. (2017). Sublingual immunotherapy for peanut allergy and induction of tolerance. *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US)*.

[27] Goetghebeur, E. and Ryan, L. (2000). Semiparametric regression analysis of interval-censored data. *Biometrics*, 56(4):1139–1144.

[28] Goggins, W. B., Finkelstein, D. M., Schoenfeld, D. A., and Zaslavsky, A. M. (1998). A markov chain monte carlo em algorithm for analyzing interval-censored data under the cox proportional hazards model. *Biometrics*, pages 1498–1507.

[29] Golightly, A. and Wilkinson, D. J. (2008). Bayesian inference for nonlinear multivariate diffusion models observed with error. *Computational Statistics & Data Analysis*, 52(3):1674–1693.

[30] Golightly, A. and Wilkinson, D. J. (2011). Bayesian parameter inference for stochastic biochemical network models using particle markov chain monte carlo. *Interface focus*, 1(6):807–820.

[31] Gómez, G., Calle, M. L., Oller, R., and Langohr, K. (2009). Tutorial on methods for interval-censored data and their implementation in r. *Statistical Modelling*, 9(4):259–297.

[32] Groeneboom, P. and Wellner, J. A. (1992). *Information bounds and nonparametric maximum likelihood estimation*, volume 19. Springer Science & Business Media.

[33] Hefle, S. L., Furlong, T. J., Niemann, L., Lemon-Mule, H., Sicherer, S., and Taylor, S. L. (2007). Consumer attitudes and risks associated with packaged foods having advisory labeling regarding the presence of peanuts. *Journal of Allergy and Clinical Immunology*, 120(1):171–176.

[34] Huang, J. et al. (1996). Efficient estimation for the proportional hazards model with interval censoring. *The Annals of Statistics*, 24(2):540–568.

[35] Huang, J. and Rossini, A. (1997). Sieve estimation for the proportional-odds failure-time regression model with interval censoring. *Journal of the American Statistical Association*, 92(439):960–967.

[36] Huang, J. and Wellner, J. A. (1997). Interval censored survival data: a review of recent progress. *LECTURE NOTES IN STATISTICS-NEW YORK-SPRINGER VERLAG-*, pages 123–170.

[37] Husain, Z. and Schwartz, R. A. (2012). Peanut allergy: an increasingly common life-threatening disorder. *Journal of the American Academy of Dermatology*, 66(1):136–143.

[38] J, S. (2006). *The Statistical Analysis of Interval-Censored Data*. Berlin: Springer.

[39] Jelliffe, R., Schumitzky, A., and Van Guilder, M. (2000). Population pharmacokinetics/pharmacodynamics modeling: parametric and nonparametric methods. *Therapeutic drug monitoring*, 22(3):354–365.

[40] Kale, B. and Muralidharan, K. (2002). Optimal estimating equations in mixture distributions accommodating instantaneous or early failures. *Quality Control and Applied Statistics*, 47(6):677–680.

[41] Kim, E. H., Bird, J. A., Kulis, M., Laubach, S., Pons, L., Shreffler, W., Steele, P., Kamilaris, J., Vickery, B., and Burks, A. W. (2011). Sublingual immunotherapy for peanut allergy: clinical and immunologic evidence of desensitization. *Journal of Allergy and Clinical Immunology*, 127(3):640–646.

[42] Knopik, L. (2011). Model for instantaneous failures. *Scientific Problems of Machines Operation and Maintenance*, 46(2):37–45.

[43] Kou, S., Olding, B. P., Lysy, M., and Liu, J. S. (2012). A multiresolution method for parameter estimation of diffusion processes. *Journal of the American Statistical Association*, 107(500):1558–1574.

[44] Lamborn, K. R., Yung, W. A., Chang, S. M., Wen, P. Y., Cloughesy, T. F., DeAngelis, L. M., Robins, H. I., Lieberman, F. S., Fine, H. A., Fink, K. L., et al. (2008). Progression-free survival: an important end point in evaluating therapy for recurrent high-grade gliomas. *Neuro-oncology*, 10(2):162–170.

[45] Langevin, P. (1908). Sur la théorie du mouvement brownien. *CR Acad. Sci. Paris*, 146(530-533):530.

[46] Lee, B. L., Kosorok, M. R., and Fine, J. P. (2005). The profile sampler. *Journal of the American Statistical Association*, 100(471):960–969.

[47] Li, J. and Ma, S. (2013). *Survival analysis in medicine and genetics*. CRC Press.

[48] Lin, X. and Wang, L. (2010). A semiparametric probit model for case 2 interval-censored failure time data. *Statistics in medicine*, 29(9):972–981.

[49] Littell, R. C. (1996). *SAS*. Wiley Online Library.

[50] Liu, C., Yang, W., Devidas, M., Cheng, C., Pei, D., Smith, C., Carroll, W. L., Raetz, E. A., Bowman, W. P., Larsen, E. C., et al. (2016). Clinical and genetic risk factors for acute pancreatitis in patients with acute lymphoblastic leukemia. *Journal of Clinical Oncology*, 34(18):2133–2140.

[51] Liu, H. and Shen, Y. (2009). A semiparametric regression cure model for interval-censored data. *Journal of the American Statistical Association*, 104(487):1168–1178.

[M et al.] M, C., Kim EH, Withana Gamage PW, M. C., and MR, K. *Study design with staggered sampling times for evaluating sustained unresponsiveness to peanut sublingual immunotherapy*. Submitted (2017) and in review. Included as Chapter 4.

[52] Matsuzaki, A., Nagatoshi, Y., Inada, H., Nakayama, H., Yanai, F., Ayukawa, H., Kawakami, K., Moritake, H., Suminoe, A., and Okamura, J. (2005). Prognostic factors for relapsed childhood acute lymphoblastic leukemia: Impact of allogeneic stem cell transplantationa report from the kyushu-yamaguchi children's cancer study group. *Pediatric blood & cancer*, 45(2):111–120.

[53] McMahan, C. S., Wang, L., and Tebbs, J. M. (2013). Regression analysis for current status data using the em algorithm. *Statistics in medicine*, 32(25):4452–4466.

[54] Muralidharan, K. (1999). Tests for the mixing proportion in the mixture of a degene-rate and exponential distribution. *Journal Indian Statistical Associations*, 37:105–119.

[55] Muralidharan, K. and Lathika, P. (2006). Analysis of instantaneous and early failures in weibull distribution. *Metrika*, 64(3):305–316.

[56] Murphy, S. A. and Van der Vaart, A. W. (2000). On profile likelihood. *Journal of the American Statistical Association*, 95(450):449–465.

[57] Murthy, D. P., Xie, M., and Jiang, R. (2004). *Weibull models*, volume 505. John Wiley & Sons.

[58] Odell, P. M., Anderson, K. M., and D'Agostino, R. B. (1992). Maximum likelihood estimation for interval-censored data using a weibull-based accelerated failure time model. *Biometrics*, pages 951–959.

[59] Overgaard, R. V., Jonsson, N., Tornøe, C. W., and Madsen, H. (2005). Non-linear mixed-effects models with stochastic differential equations: implementation of an estimation algorithm. *Journal of pharmacokinetics and pharmacodynamics*, 32(1):85–107.

[60] Pan, W. (1999). Extending the iterative convex minorant algorithm to the cox model for interval-censored data. *Journal of Computational and Graphical Statistics*, 8(1):109–120.

[61] Pan, W. (2000). A multiple imputation approach to cox regression with interval-censored data. *Biometrics*, 56(1):199–203.

[62] Pham, H. and Lai, C.-D. (2007). On recent generalizations of the weibull distribution. *IEEE transactions on reliability*, 56(3):454–458.

[63] Picchini, U., GAETANO, A. D., and Ditlevsen, S. (2010). Stochastic differential mixed-effects models. *Scandinavian Journal of statistics*, 37(1):67–90.

[64] Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical science*, pages 425–441.

[65] Roberts, G. O. and Stramer, O. (2001). On inference for partially observed nonlinear diffusion models using the metropolis–hastings algorithm. *Biometrika*, 88(3):603–621.

[66] Rosen, J. B. (1960). The gradient projection method for nonlinear programming. part i. linear constraints. *Journal of the society for industrial and applied mathematics*, 8(1):181–217.

[67] Rücker, G. and Messerer, D. (1988). Remission duration: An example of interval-censored observations. *Statistics in Medicine*, 7(11):1139–1145.

[68] Sampson, H. A., Van Wijk, R. G., Bindslev-Jensen, C., Sicherer, S., Teuber, S. S., Burks, A. W., Dubois, A. E., Beyer, K., Eigenmann, P. A., Spergel, J. M., et al. (2012). Standardizing double-blind, placebo-controlled oral food challenges: American academy of allergy, asthma & immunology–european academy of allergy and clinical immunology practall consensus report. *Journal of Allergy and Clinical Immunology*, 130(6):1260–1274.

[69] Satten, G. A. (1996). Rank-based inference in the proportional hazards model for interval censored data. *Biometrika*, 83(2):355–370.

[70] Sermaidis, G., Papaspiliopoulos, O., Roberts, G. O., Beskos, A., and Fearnhead, P. (2013). Markov chain monte carlo for exact inference for diffusions. *Scandinavian Journal of Statistics*, 40(2):294–321.

[71] Sicherer, S. H., Muñoz-Furlong, A., Godbold, J. H., and Sampson, H. A. (2010). Us prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. *Journal of Allergy and Clinical Immunology*, 125(6):1322–1326.

[72] Sicherer, S. H. and Sampson, H. A. (2014). Food allergy: epidemiology, pathogenesis, diagnosis, and treatment. *Journal of Allergy and Clinical Immunology*, 133(2):291–307.

[73] Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2005). Winbugs user manual. mrc biostatistics unit, cambridge.

[74] Stramer, O., Bognar, M., et al. (2011). Bayesian inference for irreducible diffusion processes using the pseudo-marginal approach. *Bayesian Analysis*, 6(2):231–258.

[75] Stramer, O., Bognar, M., and Schneider, P. (2009). Bayesian inference for discretely sampled markov processes with closed-form likelihood expansions. *Journal of Financial Econometrics*, 8(4):450–480.

[76] Sun, J., Hui, X., Ying, W., Liu, D., and Wang, X. (2014). Efficacy of allergen-specific immunotherapy for peanut allergy: a meta-analysis of randomized controlled trials. In *Allergy and asthma proceedings*, volume 35(2), pages 171–177. OceanSide Publications, Inc.

[77] Taylor, S. L., Crevel, R. W., Sheffield, D., Kabourek, J., and Baumert, J. (2009). Threshold dose for peanut: risk characterization based upon published results from challenges of peanut-allergic individuals. *Food and Chemical Toxicology*, 47(6):1198–1204.

[78] Taylor, S. L., Moneret-Vautrin, D., Crevel, R. W., Sheffield, D., Morisset, M., Dumont, P., Remington, B. C., and Baumert, J. L. (2010). Threshold dose for peanut: risk characterization based upon diagnostic oral challenge of a series of 286 peanut-allergic individuals. *Food and chemical toxicology*, 48(3):814–819.

[79] Tornøe, C. W., Overgaard, R. V., Agersø, H., Nielsen, H. A., Madsen, H., and Jonsson, E. N. (2005). Stochastic differential equations in nonmem®: implementation, application, and comparison with ordinary differential equations. *Pharmaceutical research*, 22(8):1247–1258.

[80] Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 290–295.

[81] Uhlenbeck, G. E. and Ornstein, L. S. (1930). On the theory of the brownian motion. *Physical review*, 36(5):823.

[82] Vickery, B. P., Berglund, J. P., Burk, C. M., Fine, J. P., Kim, E. H., Kim, J. I., Keet, C. A., Kulis, M., Orgel, K. G., Guo, R., et al. (2017). Early oral immunotherapy in peanut-allergic preschool children is safe and highly effective. *Journal of Allergy and Clinical Immunology*, 139(1):173–181.

[83] Vickery, B. P., Scurlock, A. M., Kulis, M., Steele, P. H., Kamilaris, J., Berglund, J. P., Burk, C., Hiegel, A., Carlisle, S., Christie, L., et al. (2014). Sustained unresponsiveness to peanut in subjects who have completed peanut oral immunotherapy. *Journal of Allergy and Clinical Immunology*, 133(2):468–475.

[84] Vonesh, E. and Chinchilli, V. M. (1996). *Linear and nonlinear models for the analysis of repeated measurements*. CRC press.

[85] Wang, L. and Dunson, D. B. (2011). Semiparametric bayes' proportional odds models for current status data with underreporting. *Biometrics*, 67(3):1111–1118.

[86] Wang, L., McMahan, C. S., Hudgens, M. G., and Qureshi, Z. P. (2016). A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data. *Biometrics*, 72(1):222–231.

[87] Wang, N., Wang, L., and McMahan, C. S. (2015). Regression analysis of bivariate current status data under the gamma-frailty proportional hazards model using the em algorithm. *Computational Statistics & Data Analysis*, 83:140–150.

[88] Whitaker, G. A., Golightly, A., Boys, R. J., Sherlock, C., et al. (2017). Bayesian inference for diffusion-driven mixed-effects models. *Bayesian Analysis*, 12(2):435–463.

[89] Withana Gamage PW, Chaudhari M, M. C. and MR, K. (2017). A proportional hazards model for interval-censored subject to instantaneous failures. In Submission (2017).

[90] Yu, B. (2010). A bayesian mcmc approach to survival analysis with doubly-censored data. *Computational statistics & data analysis*, 54(8):1921–1929.

[91] Yunginger, J. W., Sweeney, K. G., Sturner, W. Q., Giannandrea, L. A., Teigland, J. D., Bray, M., Benson, P. A., York, J. A., Biedrzycki, L., Squillace, D. L., et al. (1988). Fatal food-induced anaphylaxis. *Jama*, 260(10):1450–1452.

[92] Zhang, Y., Hua, L., and Huang, J. (2010). A spline-based semiparametric maximum likelihood estimation method for the cox model with interval-censored data. *Scandinavian Journal of Statistics*, 37(2):338–354.

[93] Zhang, Y. and Jamshidian, M. (2004). On algorithms for the nonparametric maximum likelihood estimator of the failure function with censored data. *Journal of Computational and Graphical Statistics*, 13(1):123–140.

[94] Zhang, Z. and Sun, J. (2010). Interval censoring. *Statistical Methods in Medical Research*, 19(1):53–70.