

RESEARCH ARTICLE

Study design with staggered sampling times for evaluating sustained unresponsiveness to peanut sublingual immunotherapy

Monica Chaudhari¹  | Edwin H. Kim² | Prabhashi W. Withana Gamage³ | Christopher S. McMahan³  | Michael R. Kosorok¹

¹Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

²Division of Allergy and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

³Department of Mathematical Sciences, Clemson University, Clemson, South Carolina

Correspondence

Monica Chaudhari, PO Box 4241, Bellevue, WA 98009, USA.
Email: monica.chaudhari@gmail.com

Funding information

National Institutes of Health, Grant/Award Number: AI121351 and UL1 TR001111

In this work, we delineate an altered study design of a pre-existing clinical trial that is currently being implemented in 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 the effectiveness of sublingual immunotherapy in achieving sustained unresponsiveness (SU) as assessed by repeated double-blind placebo-controlled food challenges (DBPCFC). With scarce published literature characterizing SU, the length of time off-therapy that would represent clinically meaningful benefit remains undefined. We use the new design features to assess time to loss of SU, an important efficacy endpoint, that to our knowledge, no prior study has investigated. Our work has two-fold objectives: first is to propose and discuss aspects of the altered design that would allow us to study SU and second is to explore methodology to evaluate the time to loss of SU and its association with risk factors in the context of the data originating from the trial. The salient feature of the new design is the allocation scheme of study subjects to staggered sampling timepoints when a subsequent DBPCFC is administered. Due to this feature, the time to loss of SU is either left or right censored. Additionally, some participants at study entry fail the DBPCFC, leading to what can be construed as an instantaneous failure. Through in-depth numerical studies, we examine the performance and power of a recently proposed mixture proportional hazards model specifically designed for the analysis of interval-censored data subject to instantaneous failures.

KEYWORDS

food allergy, instantaneous events, interval-censored data, proportional hazards model

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.^{1,2} In general, allergic reactions to peanuts may manifest minor symptoms like itchiness, urticaria, swelling, eczema, sneezing, abdominal pain, and 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 immunotherapies that primarily differ in their mechanisms of delivering small allergen doses to a patient. For example, subcutaneous immunotherapy (SCIT) exposes patients to the allergen through injections, oral immunotherapy (OIT) through direct ingestion, and sublingual immunotherapy (SLIT) through its administration under the tongue. For the purposes of desensitizing patients to food allergies, SCIT has shown little promise due to its association with high rate of systematic reactions, whereas OIT and SLIT have shown promise but require further investigations. While both are demonstrated to be effective in achieving clinical desensitization, ie, an increase in reaction threshold while receiving the study drug,³ SLIT is often favored over OIT as it is considered safer due to fewer systemic reactions and has an easier mode of administration. On the other hand, few studies have been successful in demonstrating sustained unresponsiveness (SU), ie, a persisting non-reactive state after discontinuation of therapy, which, in contrast to the phenomenon of tolerance (TLC), wanes after a period.⁴

Various studies of OIT and SLIT have looked at SU,⁵⁻⁷ but these studies either have been small or have failed to systematically investigate the loss of SU within a single population. The failure, in part, is due to the study designs that require desensitized subjects to take subsequent challenges at a specific time that could either be too early to demonstrate SU or late enough to run the risk of high subject withdrawal rate. Consequently, study designs with longer off-therapy duration 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. Through an ongoing clinical investigation of peanut SLIT in a pediatric cohort, we presented with a unique opportunity to address issues pertaining to SU by supplementing this extensive longitudinal study with innovative design and advanced statistical methodology.

The remainder of this paper 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 conducted to evaluate finite sample performance of the proposed procedure and ends in clinical trial data application and analysis. Section 5 concludes with a summary discussion.

2 | STUDY DESIGN: TRADITIONAL VS REVISED

Supported by the National Institute of Allergy and Infectious Diseases, the Sublingual Immunotherapy for Peanut Allergy and Induction of Tolerance (SLIT-TLC) and clinical trial ID is: NCT01373242 is a revised augmentation to an ongoing prospective, double-blind, randomized, placebo-controlled phase-II clinical trial that is conducted in Allergy, Immunology, and Rheumatology Division of the Department of Pediatrics at the University of North Carolina at Chapel Hill.⁸ Protocols were approved by the Human Subjects Committee at the University of North Carolina. In contrast to most studies that focus on estimating the treatment effect, the study has an objective to assess the effect of discontinuing therapy administered to the peanut-allergic subjects that received 48 months of open-label peanut SLIT.

In concurrence of both protocol versions, the study consists of a screening/baseline visit, build-up phase (approximately, 20 weeks), maintenance phase (42 months), and, lastly, TLC/SU phase. All enrolled subjects between the ages of 1 to 11 years undergo a qualifying entry double-blind placebo-controlled food challenge (DBPCFC) with 1000 mg of cumulative peanut protein (25, 50, 100, 250, 575) mg to confirm the peanut allergy diagnosis and establish a baseline threshold. Following a positive DBPCFC, each participant is 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 (approximately, 20 weeks) and to maintain this dose daily over the maintenance phase (42 months). At the completion of study drug dosing (48 months), all compliant participants then undergo a second DBPCFC to determine desensitization.

2.1 | The traditional design

Following the 48-month DBPCFC of cumulative 2500 mg of peanut protein (50, 100, 200, 400, 750, 1000) mg, the original protocol mandated that participants who do not demonstrate a clinically significant desensitized threshold (greater than or equal to 1000 mg) stop peanut SLIT with an option to consume a daily peanut food equivalent and those who do

(the desensitized subjects) randomize in a blinded fashion 2:1 to placebo or continued peanut SLIT for the following six months. Furthermore, to assess clinical tolerance, the placebo group complete the trial with 54-month DBPCFC of cumulative 5000 mg (250, 500, 1000, 1000, 1000, 1250) mg; the treatment group, on the other hand, continues open-label SLIT for an additional six months (60 months of dosing) followed by a 60-month DBPCFC and discontinues therapy for an additional six months before returning for a final 66-month DBPCFC. While tolerance among placebo group is assessed through the 48- and 54-month DBPCFC results, it is assessed through 60- and 66-month DBPCFC results among the treatment group. To assess the effect of peanut SLIT in inducing clinical tolerance, the primary endpoint was defined as the ratio of success rate within the two arms (placebo vs treatment) as determined by 54-month DBPCFC. To examine the rate of tolerance induction, the protocol proposed a secondary endpoint that measured success rate following both 48 months and 60 months of peanut SLIT study drug.

This design presented several challenges. (i) The 54-month challenge as designed would likely put the placebo group (two third of the participants) to undue high challenge doses when thresholds, in the absence of therapy, are expected to drop. (ii) The primary and secondary endpoints provided cross-sectional summary estimates and failed to account for within-subject changes in threshold. (iii) The outcome measures suffered from small sample size of the treatment arm and for the effect to be significant required unfeasibly large number of successes compared to the placebo arm. (iv) The long trial ran a risk of noncompliance among the pediatric cohort. (v) Last but not least, based on the results from the work of Vickery et al⁵ that showed only 50% tolerance after a month of using the presumed more robust OIT lead to suspicion that six months of SLIT would not result in success.

Recognizing that the six-month period to assess clinical tolerance was rather optimistic, arbitrary, and less supported by prior research and that the 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 the final post-avoidance (PA) DBPCFC among desensitized participants. This dissolved the need for a two-arm trial and shortened the length of the overall clinical trial. Furthermore, acknowledging that the therapy effect may not last forever upon its discontinuation, “tolerance” was redefined to reflect SU. 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 was conducted to evaluate the safety and efficacy of peanut SLIT in inducing clinical SU among peanut allergic children.

2.2 | The revised design

To assess desensitization, the 48-month DBPCFC at the end of the maintenance phase was redesigned to a cumulative 5000 mg of peanut protein and is discussed in detail in the following. The desensitization criterion for participants to enter the following SU phase of this study includes meeting the minimal clinically relevant threshold (MCRT) of 300 mg of peanut protein without symptoms.⁹ Those who fail are required to stop peanut SLIT and conclude the study, with a recommendation to resume a strict peanut avoidance diet. On the other hand, the desensitized participants enter the SU phase during which peanut SLIT is discontinued. They are block randomized to one of the 17 weeks when the PA DBPCFC is administered to assess SU. This marks study completion, and at the primary investigators clinical discretion, the participants are recommended to transition to a daily peanut food equivalent to maintain the desensitized effect.

2.2.1 | Randomized weekly assignment for the PA food challenge

Before the start of the study's 17-week SU phase, a computer-generated block randomization schedule was prepared such that the participants who complied to 48 months of dosing had equal probability of being randomized to one of the following time periods: initial 6 weeks, between 7 and 12 weeks, and between 13 and 17 weeks. A participant was then assigned to a week sampled from a six-week block that he/she was randomized to. Based on this randomization schedule, three participants were randomly assigned to the weekly administered PA during the SU phase to evaluate for the loss of SU effect. The PA DBPCFC, which has a structure similar to the 48-month and is administered to only desensitized participants, designates 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 48-month DBPCFC. At the primary investigator's clinical discretion, participants are then recommended to transition to a daily peanut food equivalent to maintain the desensitized effect.

2.2.2 | Double-blind placebo-controlled food challenge

Each desensitized participant is administered the protocol DBPCFC at baseline, at the end of 48 months of dosing 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, a participant is required to have performed a physical exam and peak expiratory flow measurements. Unlike latter challenges, the entry challenge is administered to a lower dose simply to prove allergy status and not the treatment response.

The DBPCFC consists of randomly ordered administration of two parts: one that consists of graded doses of peanut flour and the other that consists of identical graded doses of placebo in the form of oat flour. To monitor for symptoms, the doses are given every 10 to 20 minutes up to a cumulative dose. Before resuming the second part of the challenge, a participant is allowed one hour for washout of dosing and is observed for side effects from the first part of the challenge. Reactions are scored using a Food Challenge Symptom Score sheet. In case of significant objective or persistent subjective symptoms, the food challenge is terminated and the participant is given appropriate treatment. Symptomatic participants are then observed for a minimum of two hours before being discharged from the clinical research unit.

2.2.3 | The post-dosing DBPCFC structure and time boundaries

In contrast to the standard challenges designed in such studies with multiplicative dose increments,^{10,11} the post-dosing DBPCFCs were inspired by the Fibonacci series,¹² which allows the challenges to serve as a good check of observable cumulative effect of doses. The DBPCFC structure is illustrated in Figure 1. The rationale for the starting dose of 100 mg was based on the findings of the work of Kim et al¹³ that had the median tolerated dose of 1700 mg with a wide range (85 mg after a year - 500 mg after 3+ years of treatment). The challenge size was kept at cumulative 5000 mg of peanut protein to avoid false negatives,¹⁰ chosen arbitrarily within the range of what might mostly be eaten (~ 1 tbs of peanut butter or 4000 mg of peanut protein). Considering these dose constraints, the end challenge design was listed as (100, 200, 500, 800, 1300, 2100) mg with the first Fibonacci dose of 300 mg (MCRT dose level), broken into two incremental doses of 100 and 200 mg.

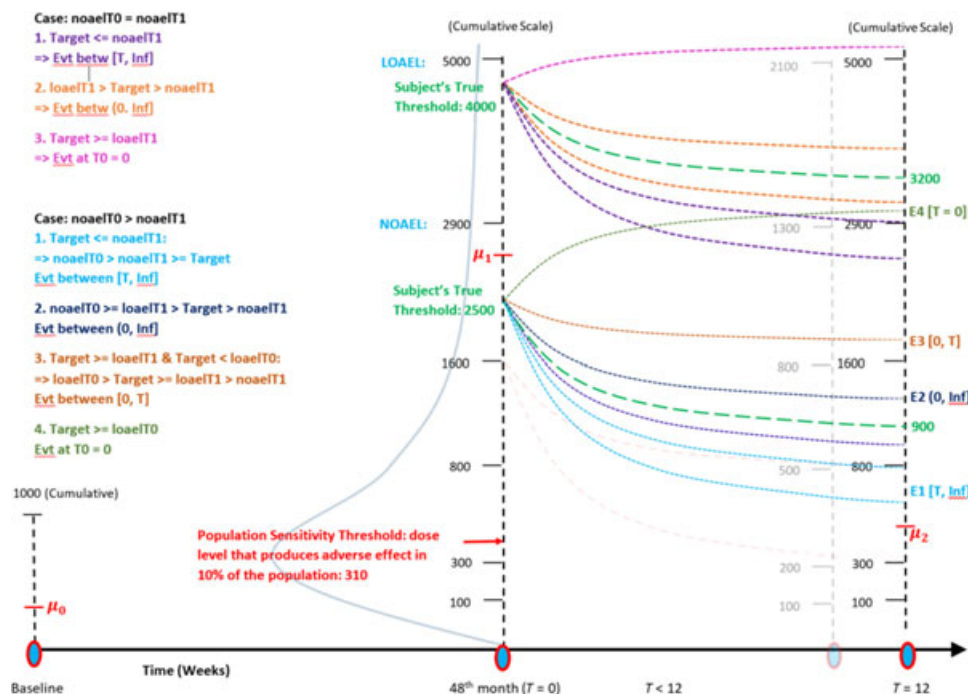


FIGURE 1 Peanut SLIT-TLC study design⁸: illustration of an exhaustive set of targeted thresholds positioned relative to a participant's observed threshold intervals and latent true threshold trajectory (the broad dashed line) for two possible scenarios: (i) when the no-observed-adverse-effect level (NOAEL) at post-avoidance (PA) double-blind placebo-controlled food challenge (DBPCFC) remains unchanged and (ii) when the NOAEL at PA DBPCFC is at least a level lower. All trajectories other than a participant's true threshold trajectory end in targeted thresholds. For example, when the true latent PA threshold is 900, a participant is expected to meet the targeted threshold of ≤ 800 at or beyond the time when the PA DBPCFC is administered [Colour figure can be viewed at wileyonlinelibrary.com]

Consequently, the design allowed not only an observable MCRT of 300 mg but also more granularity above the 500-mg dose level where most participants were expected to reach their threshold while maintaining the total maximum challenge size of 5000 mg. The gradual dose gradation of this structure mitigated risk at higher doses by providing enhanced safety from more frequent monitoring of adverse events.

Each administered DBPCFC results in subject-specific threshold bounds: no-observed-adverse-effect level (NOAEL), the highest administered cumulative dose that does not elicit any adverse event, and lowest-observed-adverse-effect level (LOAEL), the lowest that does.¹⁴ Let $D_i^{T_0} \in [N_i^{T_0}, L_i^{T_0}]$ define the i^{th} participant's true desensitization threshold achieved at the 48-month DBPCFC, ($T_0 = 0$), and $D_i^{T_{1i}} \in [N_i^{T_{1i}}, L_i^{T_{1i}}]$, the true tolerance threshold achieved at the PA DBPCFC at randomized time, T_{1i} , such that the observed bounds maintain $D_{\text{MCRT}}^{T_0} < N_i^{T_{1i}} \leq N_i^{T_0}$. Considering tolerance to a targeted dose $D_T = d$ in the absence of immunotherapy to be a nonincreasing process, a participant's time to loss of SU, T_i^d , is well approximated by an interval censoring framework as illustrated in Figure 1. An exhaustive set of targeted dose thresholds is positioned relative to a participant's observed 48-month and PA NOAELs that help determine threshold trajectories for the two possible scenarios: (i) $N_i^{T_{1i}} = N_i^{T_0}$ and (ii) $N_i^{T_{1i}} < N_i^{T_0}$. 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, ie, 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, ie, 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 1, is summarized using the following observables:

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

We note that $N_i^{T_{1i}} \leq N_i^{T_0}$ is always true under the study design construct. Furthermore, a target dose threshold that lies within an unchanged observed threshold interval does not contribute to any information on time to failure. A participant's loss of SU to an arbitrary targeted dose level, $D_T (> D_{\text{MCRT}}^{T_0})$, is earliest claimed at T_{1i} if $N_i^{T_{1i}} \geq D_T$. The nondesensitized participants with $N_i^{T_0} \leq D_{\text{MCRT}}^{T_0}$ 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 that are subject to instantaneous failures. Furthermore, it is worthwhile to note that, due to the study design, the censoring times are noninformative since the participants are administered PA DBPCFCs at randomized observation times, suggesting that the observation process is independent of the failure times of interest.

3 | MODEL AND METHODS

Herein, a brief discussion of the statistical methods is provided. Let $T^d \equiv T$ denote the survival time of interest for a pre-selected targeted dose d . The mixture model proposed in the work of Withana Gamage et al¹⁵ provides the cumulative distribution function (cdf) of T as

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

where $\mathbf{x}$ is an $(r \times 1)$ -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 proportional hazards 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 proportional hazards assumption. Second, one will note that the mixture model allows for a probability mass at time zero 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 effects 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 participant with covariates $\mathbf{x}$ has a probability of experiencing an instantaneous failure given by

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

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

$$\begin{aligned} P(T > t | \mathbf{x}) &= P(T > 0 | \mathbf{x})P(T > t | \mathbf{x}, t > 0) \\ &= e^{-\alpha e^{\alpha' \mathbf{x}}} S(t | \mathbf{x}), \end{aligned}$$

where $S(t | \mathbf{x})$ is as defined in (1).

The unknown parameters in the mixture model depicted in (1) involve the regression parameters β , 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 (1) 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, Withana Gamage et al¹⁵ suggest modeling $\Lambda_0(\cdot)$ through the use of the monotone splines in the work of Ramsay,¹⁶ ie, specify

$$\Lambda_0(t) = \sum_{l=1}^k \gamma_l b_l(t), \quad (2)$$

where $b_l(\cdot)$ are spline basis functions and γ_l are unknown spline coefficients, for $l = 1, \dots, k$. To assure that $\Lambda_0(\cdot)$ is nondecreasing, γ_l should be nonnegative, ie, $\gamma_l \geq 0$, for all $l = 1, \dots, k$. 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 on spline specification and the choice of k , see other works.¹⁵⁻¹⁸

3.1 | Estimation and inference

For participants not experiencing an instantaneous failure, let T_1 denote the time at which an individual is administered the PA 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, ie, an instantaneous failure ($T = 0$), T is left censored ($0 < T < T_1$), or T is right censored ($T > T_1$). Let ψ be the indicator denoting time to failure is beyond time 0, ie, $\psi = 1$, if T is not an instantaneous failure. Similarly, let δ be a censoring indicator taking value of 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, eg, the works of Liu and Shen¹⁹ and Zhang and Sun²⁰ and the references therein), this assumption is valid in this application by virtue of the study design that makes use of staggered randomized sampling times. The observed data collected on n subjects is given by $\mathcal{D} = \{(T_{1i}, \mathbf{x}_i, \psi_i, \delta_i); i = 1, 2, \dots, n\}$, 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

$$L(\theta) = \prod_{i=1}^n [F(T_{1i} | \mathbf{x}_i)^{\delta_i} \{1 - F(T_{1i} | \mathbf{x}_i)\}^{1-\delta_i}]^{\psi_i} \left\{ e^{-\alpha e^{\alpha' \mathbf{x}_i}} \right\}^{\psi_i} \left\{ 1 - e^{-\alpha e^{\alpha' \mathbf{x}_i}} \right\}^{1-\psi_i}, \quad (3)$$

where $\theta = (\beta', \gamma', \alpha)'$ represents the set of unknown parameters that need to be estimated, with $\gamma = (\gamma_1, \dots, \gamma_k)'$. Since the observed data likelihood exists in closed form, one could obtain the maximum likelihood estimator of θ , as $\hat{\theta} = \arg\max_{\theta} L(\theta)$, through numerically maximizing the logarithm of (3). However, due to numerical instabilities of this process,²¹ it is generally suggested that the expectation-maximization algorithm developed and detailed in the work of Withana Gamage et al¹⁵ be used for the purposes of identifying the maximum likelihood estimator, which we denote herein as $\hat{\theta} = (\hat{\beta}', \hat{\gamma}', \hat{\alpha})'$.

For uncertainty quantification, two variance estimators were considered and evaluated: the outer product of gradients (OPG) estimator¹⁵ and the Huber sandwich estimator.²² While both estimators, in general, seem to be viable, the results

from extensive numerical studies (not shown) found the Huber sandwich estimator more robust and less sensitive to convergence issues in small sample size application similar to the one described in Section 2. In general, the OPG estimator is obtained as

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

and the Huber sandwich estimator is obtained as

$$\hat{V}_{\text{HS}}(\hat{\theta}) = -A^{-1}B(-A)^{-1},$$

where $A = \sum_{i=1}^n \partial^2 l_i(\theta) / \partial \theta^2 |_{\theta=\hat{\theta}}$, $B = \sum_{i=1}^n \hat{g}_i(\hat{\theta}) \hat{g}_i'(\hat{\theta})$, $\hat{g}_i(\hat{\theta}) = \partial l_i(\theta) / \partial \theta |_{\theta=\hat{\theta}}$, and $l_i(\theta)$ is the log-likelihood contribution of the i^{th} 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.

4 | NUMERICAL EXPERIMENTS

We performed the following simulation studies to investigate the finite sample performance of the proposed procedure within the context of the study described in Section 2.

4.1 | Simulation study I

The true cdf of the failure times was specified as

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

where $\mathbf{x} = (x_1, x_2)'$, $x_1 \sim N(0, 1)$, $x_2 \sim \text{Bernoulli}(0.5)$, $p_0 \in \{0.1, 0.2, 0.3\}$, and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$, where β_1 and β_2 each take on values -0.5 and 0.5 leading to four regression parameter configurations. Recall that the relation between p_0 and α is given by $\alpha = -\log(1 - p_0)$. The true cumulative baseline hazard function was specified as $\Lambda_0(t) = 0.1t$. The choice of an exponential distribution for the 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 datasets, each consisting of individual level data including failure times, T , for n subjects, where $n \in \{50, 100, 200, 400\}$. The observation times, T_1 , were generated to reflect those employed in the study (see Section 2.2.1). The left- or right-censoring indicators for the failure times were determined based on whether the events occurred earlier or later than T_1 , respectively.

The cumulative baseline hazard function was specified using the monotone spline representation outlined in Section 3 with basis functions of degree 2, one interior knot placed at the median, and boundary knots placed at the minimum and maximum of the observation times. Finally, the expectation-maximization algorithm¹⁵ was used to fit the model and convergence was declared when the maximum absolute difference of all of the consecutive parameter updates was less than 1×10^{-5} . To estimate the standard errors for the estimated regression coefficients, we used the aforementioned Huber sandwich estimator. This procedure was repeated for each dataset.

Table 1 provides a summary of the estimated regression coefficients across all of the considered parameter configurations for each sample size, $n \in \{50, 100, 200, 400\}$. This summary includes the sample mean, the mean standard error estimates, sample standard deviation of the 500 estimates, and power to detect the effects 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 is generally in agreement with the mean standard error estimates. This suggests that the Huber sandwich estimator performs reasonably well in estimating asymptotic variances. To estimate power, a Wald-type inference was conducted using the sandwich estimator to test the set of hypotheses $H_0 : \beta_j = 0$ versus $H_1 : \beta_j \neq 0$ for each of the 500 datasets at 0.05 significance level and the proportion of rejections was reported. As one might expect, the empirical power increases with the sample size. Furthermore, the power to detect a significant effect associated with a continuous covariate dominates the power associated with a dichotomous covariate, which again is expected. Figure 2 illustrates these findings for the case when $p_0 = 0.2$.

TABLE 1 Simulation study I: reported results summarize the regression parameter estimates across all considered configurations (varying n , p_0 , β), each with 500 simulated datasets. This summary includes the average point estimates (Est), the average of the estimated standard errors (SE), the standard deviation (SD), and empirical power (Power) to detect the effect at 0.05 significance level

p_0	Parameter	$n = 50$				$n = 100$				$n = 200$				$n = 400$			
		Est	SE	SD	Power	Est	SE	SD	Power	Est	SE	SD	Power	Est	SE	SD	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

Figure 3 summarizes the estimated baseline (ie, when $\mathbf{x} = \mathbf{0}_r$) survival functions for the four 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 pointwise estimates (dashed lines), and the 2.5th and the 97.5th pointwise quantiles of the estimates (dotted lines). From these curves, one will note that the estimated baseline survival functions and, hence, the cumulative baseline hazard functions are accurately estimated and the proposed procedure can accurately estimate p_0 , which is evidenced by the difference between 1 and the estimated baseline survival functions at time zero.

4.2 | Simulation study II

This experiment was designed to emulate the data expected from the clinical trial outlined in Section 2. The data generating procedure was identical to that described in the previous study, with a few minor alterations. In particular, we set $n = 50$ and $p_0 = 0.2$ that respectively reflect an approximate sample size of the clinical trial and the expected rate of instantaneous failure (ie, the proportion of participants that fail to be desensitized by the 48-month DBPCFC). The primary interest in this experiment was to assess the power of the procedure to test 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. Thus, to examine the procedure's power to detect significant effects ranging from a small to large signal level, we specified $\beta_1 \in \{0, 0.5, 1\}$ and $\beta_2 \in \{-1, -0.5, 0, 0.5, 1\}$ for the continuous and dichotomous covariate, respectively. Negative values of β_1 were not considered due to the symmetry of continuous covariate coefficient estimate about zero. Since the symmetry does not hold for a dichotomous covariate, we included negative values of β_2 . Table 2 provides summary similar to that offered in Table 1 but for the parameter estimates arising from this experiment.

The results herein reinforce the main findings discussed in the previous simulation study. Figure 4 illustrates that the procedure will allow us to reliably detect important effects with a sample size of 50, as long as they are of a moderate size

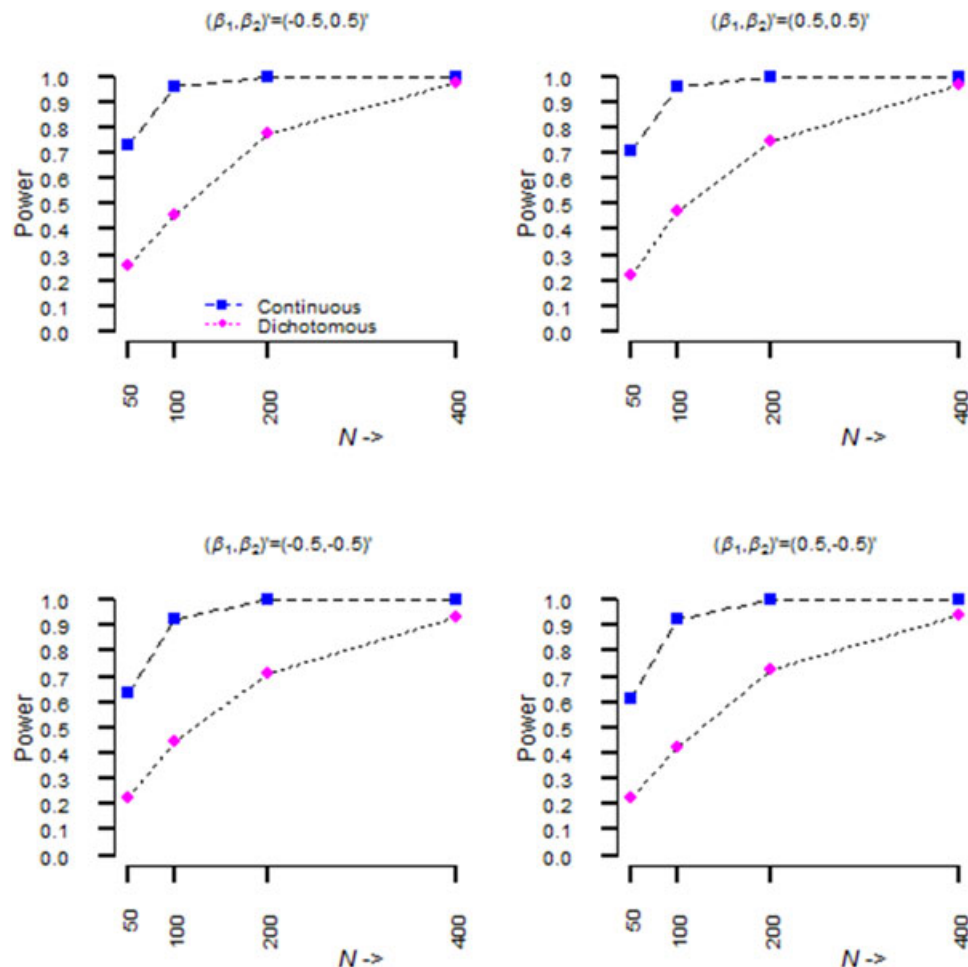


FIGURE 2 Simulation study I: these figures provide empirical power curves, as a function of sample size n , for the different configurations of the regression parameters. Note that the dashed and dotted lines represent empirical power curves for the effects associated with the continuous and dichotomous covariates, respectively [Colour figure can be viewed at wileyonlinelibrary.com]

and that the size of the test is reasonably preserved. It is worthwhile to point out that we do not assess the power of the procedure for negative effect sizes of the continuous covariate due to symmetry.

In summary, from the findings of these experiments, we conjecture that the proposed methodology is a reliable statistical approach that can be used to analyze the clinical trial data.

4.3 | Data application

Of the 54 study enrollees that successfully cleared baseline challenge, seven withdrew from the study prior to completion of SLIT dosing (48 months) and one is continuing to dose. Thus, the analyses presented is based on the cohort of 46 participants of whom two did not meet MCRT (300 mg) at the 48-month DBPCFC, four failed to react to the dose equivalent to the 48-month administered LOAEL (these subjects are considered to have PA threshold interval similar to that exhibited in 48-month DBPCFC), and lastly, seven subjects withdrew prior to their PA challenge (for conservative treatment effect, these subjects are considered to have lost tolerance at the 48-month challenge and, hence, instantaneous failures for analysis purposes).

Based on 46 participants, we performed three sets of analyses corresponding to the investigator selected target thresholds that were deemed clinically relevant: 300 mg (MCRT), 1000 mg (baseline threshold considered for eligibility), and 2000 mg (approximate median threshold evaluated at the 48-month DBPCFC). In respective analyses, the participants had the following number of events met: instantaneous (9, 10, 20), left censored (0, 11, 9), and right censored (37, 25, 17). Dichotomous predictors included gender, history of Asthma, and history of Atopic Dermatitis (AD); continuous predictors included age at consent (years) centered on zero, baseline clinical biomarkers (IgE (kU/L), IgG4 (mg/L), SPT (mm)),

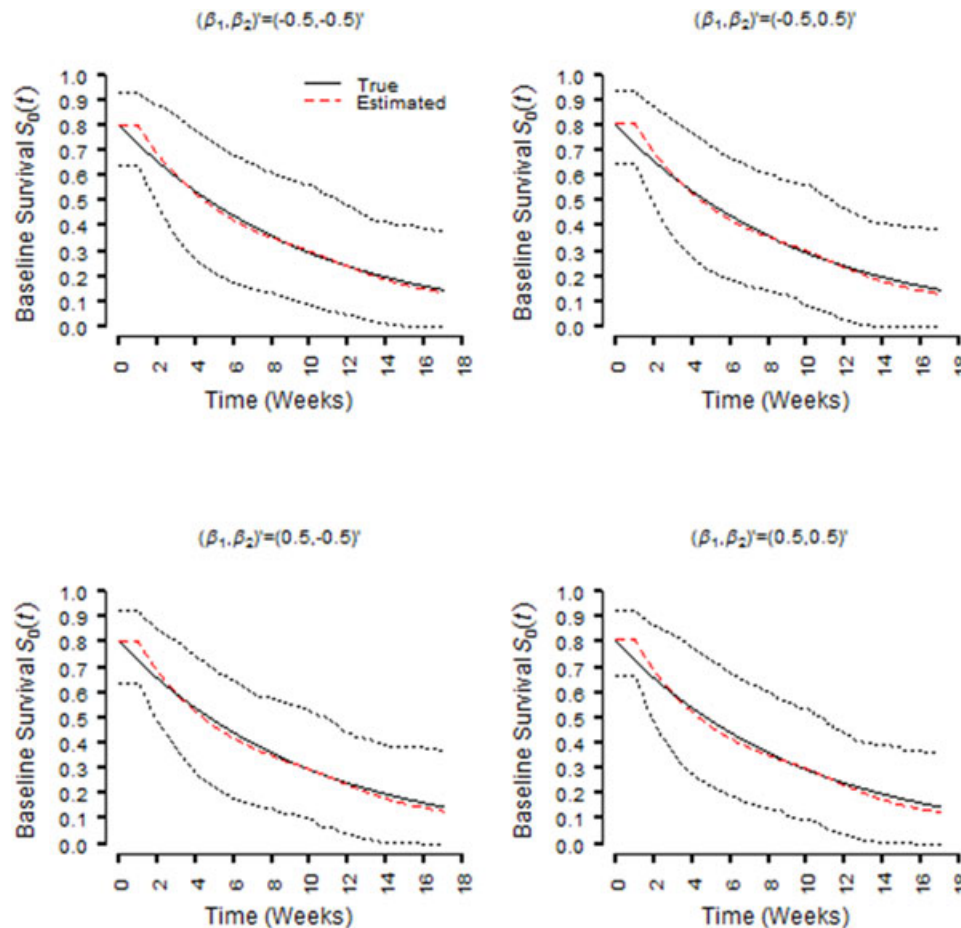


FIGURE 3 Simulation study I: these figures depict survival functions for the four different configurations of the regression parameters when $n = 50$ and $p_0 = 0.2$. Provided results include the true baseline survival function (solid lines), the average of the pointwise estimates (dashed lines), and the 2.5th and 97.5th pointwise quantiles of the estimates (dotted lines) [Colour figure can be viewed at wileyonlinelibrary.com]

and respective change from baseline at the 48th month (IgEdrop, IgG4raise, and SPTdrop) on log scale. The cumulative baseline hazard function was specified using the monotone spline representation outlined in Section 3 with basis functions of degree 3, two interior knots, and boundary knots placed at the minimum and maximum of the observation times.

Table 3 provides a summary of the estimated regression coefficients corresponding to the selected target thresholds. In the adjusted analyses, higher baseline IgE and lower baseline IgG4 and SPT are associated with increased hazard where as higher percentage drop in IgE is associated with a reduced hazard of the loss of SU. While significance at 0.1 level may not be achieved for all the three target thresholds (limiting to two or more), the biomarker effects do strengthen with rising targets. Additionally, the male cohort (vs female) presents lower hazard of the loss of SU; however, this difference ceases at higher target thresholds. These findings are consistent with exploratory analyses and scientific rationale. Figure 5 demonstrates gender and AD differences in probability estimates of survival to a target threshold of 1000 mg among the nonasthmatic group at average age and clinical biomarkers. Additionally, Figure 6 illustrates probability estimates of survival to varying levels of target thresholds and how they differ by gender among the nonasthmatic group with AD at average age and clinical biomarkers.

5 | DISCUSSION

The clinical design of this study emanates from a revised protocol of an ongoing 66-month SLIT study. The original study randomized desensitized participants at 48th month in a blinded fashion (2:1) to a placebo or treatment group, with the treatment group continuing peanut SLIT for six months. To assess clinical tolerance, the original protocol mandated that

TABLE 2 Simulation study II: reported results summarize the regression parameter estimates across all considered configurations ($n = 50$, $p_0 = 0.2$, varying β) each with 500 simulated datasets. This summary includes the average point estimates (Est), the average of the estimated standard errors (SE), the standard deviation (SD), and empirical power (Power) to detect the effect at 0.05 significance level

Parameter	Est	SE	SD	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

the placebo group complete the trial with a 54-month DBPCFC, whereas the treatment group complete the trial with a 66-month DBPCFC. Due to concerns about negligible success rates, the ongoing study was revised in hopes of effectively assessing clinical SU within participants after SLIT discontinuation. Since time to loss of SU 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 multiplicative dose increments of a traditional design. In contrast to traditionally larger gaps at higher dose levels, the proposed design helps mitigate participants' 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. Additionally, the new design 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, thus enabling identification of both NOAELs and LOAELs for most subjects. Narrower dose intervals also allow 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.

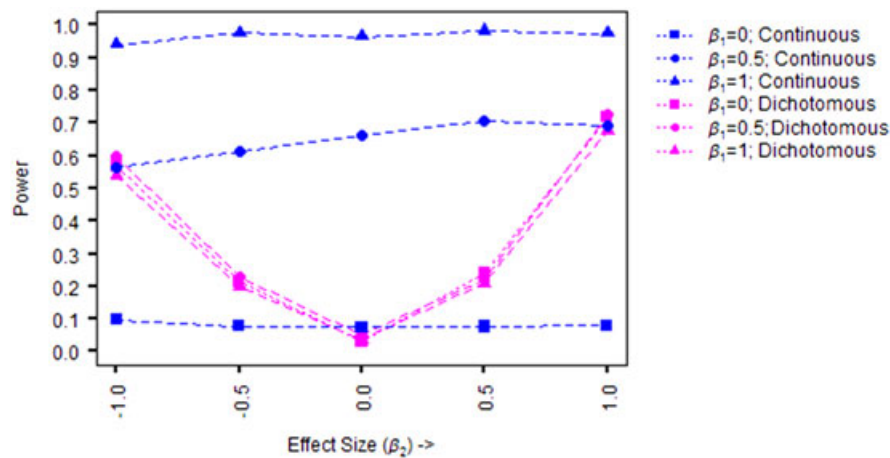


FIGURE 4 Simulation study II: this figure provides empirical power curves, as a function of the effects sizes, for the different configurations of the regression parameters [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Data application ($n = 46$): reported results summarize the parameter estimates and 95% confidence intervals for selected target thresholds: (300, 1000, 2000 mg). Baseline clinical variables (IgE, IgG4, SPT) are on the log scale. “Change from baseline” corresponds to the multiplicative change in clinical variables from the baseline

		$T_D = 300$			$T_D = 1000$			$T_D = 2000$		
		Est	SE	p-value	Est	SE	p-value	Est	SE	p-value
Demographic covariates										
	Mean centered age	-0.05	0.23	0.835	-0.18	0.12	0.153	0.01	0.24	0.478
	Male	-1.26	0.98	0.197	-0.81	0.44	0.066	-0.38	0.56	0.246
	History of asthma	-0.20	0.97	0.836	0.24	0.70	0.736	-0.34	0.70	0.313
	History of AD	-0.20	0.88	0.822	-0.72	0.49	0.142	-0.30	1.29	0.407
Clinical covariates										
Baseline										
	IgE	0.49	0.36	0.173	0.44	0.17	0.011	0.50	0.29	0.044
	IgG4	-0.39	0.33	0.234	-0.38	0.23	0.094	-0.49	0.22	0.012
	SPT	-1.54	0.76	0.042	-1.47	0.47	0.002	-0.16	0.67	0.408
Change from baseline at 48th month										
	IgE drop	-0.36	0.31	0.242	-0.36	0.19	0.065	-0.62	0.44	0.081
	IgG4 raise	0.13	0.20	0.531	0.08	0.15	0.618	0.26	0.33	0.212
	SPT drop	0.17	0.55	0.756	0.12	0.37	0.741	0.15	1.04	0.444

Abbreviations: AD, Atopic Dermatitis; SE, standard error.

The occurrence of events at study entry (ie, 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 risk set 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, ie, 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 MCRT 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 the work of Withana Gamage et al,¹⁵ for the analysis of interval-censored data subject to instantaneous failures. Through extensive simulation studies, this new methodology is shown to provide an accurate and efficient analysis in this context.

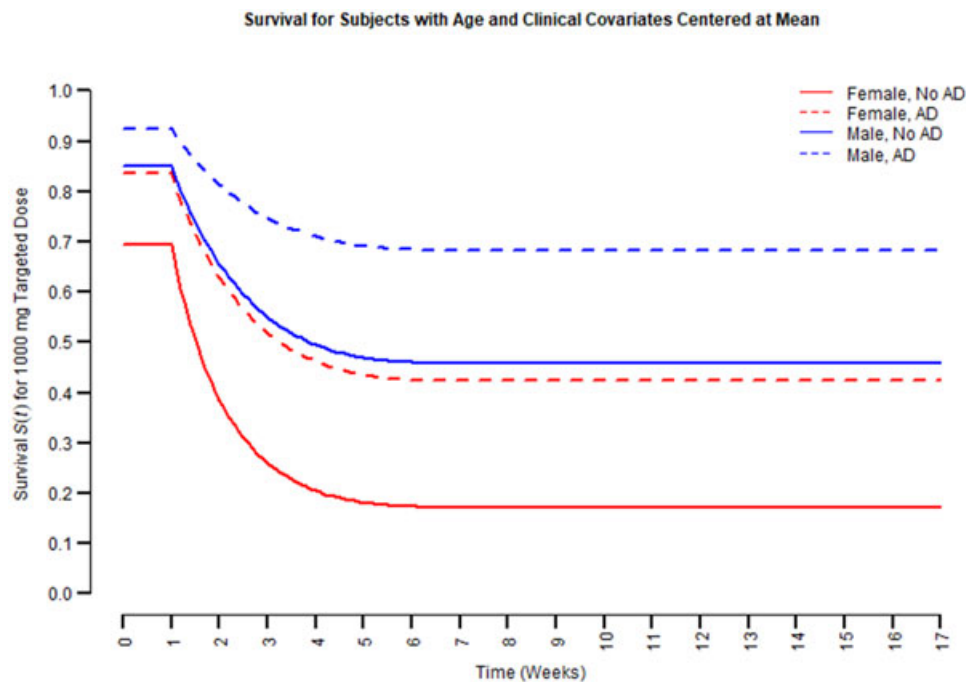


FIGURE 5 Data application: this figure provides survival curves to a targeted threshold of 1000 mg calibrated at average age and clinical biomarkers for the nonasthmatic group. AD, Atopic Dermatitis [Colour figure can be viewed at wileyonlinelibrary.com]

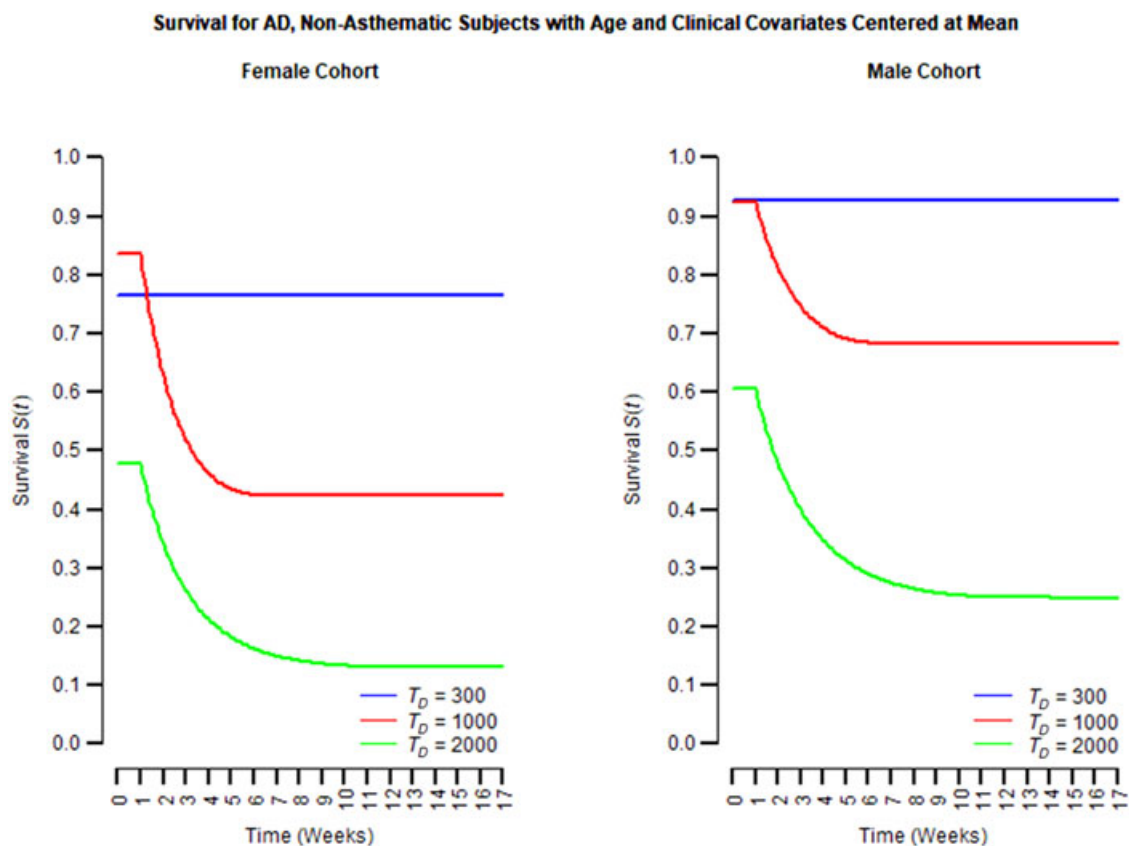


FIGURE 6 Data application: this figure provides survival curves to varying levels of target threshold: 300, 1000, and 2000 mg calibrated at average age and clinical biomarkers for the nonasthmatic group with AD [Colour figure can be viewed at wileyonlinelibrary.com]

The simulation studies presented herein illustrate the finite sample performance of the methodology developed in the aforementioned work¹⁵ 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 (1), ie, the regression coefficients, the survival functions, and the baseline probability of experiencing an instantaneous failure. Furthermore, the inferential techniques considered tend to preserve the specified size of the test, when the null is in fact true, and exhibit sufficient power to detect departures from the null for both continuous and dichotomous covariates. Through the numerical studies conducted herein and data application, we show that the methodology developed in the aforementioned work¹⁵ is capable of detecting clinically important effects in the data collected from the clinical trial.

In summary, we have illustrated a study design and accompanying method of analysis for interval-censored outcomes subject to instantaneous failures. The staggered sampling examination times of PA challenge allows the observational process to be independent of the failure times, which is required by the methodology under study and many previous proposals. The revised protocol 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 15 months. This significantly reduced the chance of noncompliance and drop-outs. Furthermore, with a few changes in the study design and efficacy analysis plan, the revised protocol allowed us to investigate new objectives of inquiry. We are hopeful that the knowledge shared here will help facilitate future studies in this therapeutic area.

ACKNOWLEDGEMENT

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

ORCID

Monica Chaudhari  <http://orcid.org/0000-0002-0854-1843>

Christopher S. McMahan  <http://orcid.org/0000-0001-5056-9615>

REFERENCES

1. Husain Z, Schwartz RA. Peanut allergy: an increasingly common life-threatening disorder. *J Am Acad Dermatol*. 2012;66(1):136-143. <https://doi.org/10.1016/j.jaad.2011.02.031>
2. Sicherer SH, Muñoz-Furlong A, Godbold JH, Sampson HA. US prevalence of self-reported peanut, tree nut, and sesame allergy: 11-year follow-up. *J Allergy Clin Immunol*. 2010;125(6):1322-1326.
3. Sun J, Hui X, Ying W, Liu D, Wang X. Efficacy of allergen-specific immunotherapy for peanut allergy: a meta-analysis of randomized controlled trials. *Allergy Asthma Proc*. 2014;35(2):171-177.
4. Burks AW, Jones SM, Wood RA, et al. Oral immunotherapy for treatment of egg allergy in children. *N Engl J Med*. 2012;367(3):233-243.
5. Vickery BP, Scurlock AM, Kulis M, et al. Sustained unresponsiveness to peanut in subjects who have completed peanut oral immunotherapy. *J Allergy Clin Immunol*. 2013;133(2):468-475.
6. Vickery BP, Berglund JP, Burk CM, et al. Early oral immunotherapy in peanut-allergic preschool children is safe and highly effective. *J Allergy Clin Immunol*. 2017;139(1):173-181.e8.
7. Burks AW, Wood RA, Jones SM, et al. Sublingual immunotherapy for peanut allergy: long-term follow-up of a randomized multicenter trial. *J Allergy Clin Immunol*. 2015;135(5):1240-1248.
8. National Center for Complementary and Integrative Health (NCCIH). Sublingual Immunotherapy for Peanut Allergy and Induction of Tolerance. Chapel Hill, NC: University of North Carolina at Chapel Hill. ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000- [cited 2017 Mar 27]. Available from: <http://clinicaltrials.gov/show/NCT01373242> NLM Identifier: NCT01373242.
9. Baumert JL, Taylor SL, Koppelman SJ. Quantitative assessment of the safety benefits associated with increasing clinical peanut thresholds through immunotherapy. *J Allergy Clin Immunol Pract*. 2018;6(2):457-465.
10. Sampson HA, Gerth van Wijk R, Bindslev-Jensen C, et al. Standardizing double-blind, placebo-controlled oral food challenges: American Academy of Allergy, Asthma & Immunology-European Academy of Allergy and Clinical Immunology PRACTALL consensus report. *J Allergy Clin Immunol*. 2012;130(6):1260-1274. <https://doi.org/10.1016/j.jaci.2012.10.017>
11. Taylor SL, Moneret-Vautrin DA, Crevel RWR, et al. Threshold dose for peanut: risk characterization based upon diagnostic oral challenge of a series of 286 peanut-allergic individuals. *Food Chem Toxicol*. 2010;48(3):814-819. <https://doi.org/10.1016/j.fct.2009.12.013>
12. Hazewinkel M. *Encyclopaedia of Mathematics: Volume 6: Subject Index - Author Index*. Dordrecht, The Netherlands: Springer Netherlands; 2013.
13. Kim EH, Bird JA, Kulis M. Sublingual immunotherapy for peanut allergy: clinical and immunologic evidence of desensitization. *J Allergy Clin Immunol*. 2011;127(3):640-646.e1. <https://doi.org/10.1016/j.jaci.2010.12.1083>
14. Taylor SL, Crevel RWR, Sheffield D, Kabourek J, Baumert J. Threshold dose for peanut: risk characterization based upon published results from challenges of peanut-allergic individuals. *Food Chem Toxicol*. 2009;47(6):1198-1204.
15. Withana Gamage PW, Chaudhari M, McMahan CS, Kosorok MR. A proportional hazards model for interval-censored data subject to instantaneous failures. arXiv:1804.00096.

16. Ramsay JO. Monotone regression splines in action. *Stat Sci.* 1988;3:425-441. <https://doi.org/10.1214/ss/1177012761>
17. Wang L, McMahan CS, Hudgens MG, Qureshi ZP. A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data. *Biometrics.* 2016;72(1):222-231. <https://doi.org/10.1111/biom.12389>
18. McMahan CS, Wang L, Tebbs JM. Regression analysis of current status data using the EM algorithm. *Statist Med.* 2013;32(25):4452-4466.
19. Liu H, Shen Y. A semiparametric regression cure model for interval-censored data. *J Am Stat Assoc.* 2009;104(487):1168-1178. <https://doi.org/10.1198/jasa.2009.tm07494>
20. Zhang Z, Sun J. Interval censoring. *Stat Methods Med Res.* 2010;19(1):53-70. <https://doi.org/10.1177/0962280209105023>
21. Wang N, Wang L, McMahan CS. Regression analysis of bivariate current status data under the Gamma-frailty proportional hazards model using the EM algorithm. *Comput Stat Data Anal.* 2015;83:140-150. <https://doi.org/10.1016/j.csda.2014.10.013>
22. Huber PJ. The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*. San Diego, CA: University of California Press; 1967:221-233.

How to cite this article: Chaudhari M, Kim EH, Withana Gamage PW, McMahan CS, Kosorok MR. Study design with staggered sampling times for evaluating sustained unresponsiveness to peanut sublingual immunotherapy. *Statistics in Medicine*. 2018;37:3944–3958. <https://doi.org/10.1002/sim.7857>