GLOBAL SENSITIVITY ANALYSIS FOR RANDOMIZED TRIALS WITH INFORMATIVE ASSESSMENT TIMES: A FULLY PARAMETRIC APPROACH

by
Zebin Wang

A thesis submitted to Johns Hopkins University in conformity with the requirements for

the degree of Master of Science

Baltimore, Maryland
April 2020

Abstract

Many randomized trials are designed to collect outcomes at fixed points in times after randomized. In practice, study participants miss their assessments or are assessed off-schedule. If and when individuals are assessed may be informative. In this thesis, we develop a sensitivity analysis methodology for analyzing randomized trials with a potentially informative assessment process. We develop these methods in the context of the *Asthma Research for the Community (ARC)* trial.

Primary Reader and Advisor: Daniel O. Scharfstein

Secondary Reader: Elizabeth A. Colantuoni

Acknowledgments

I am very grateful to my thesis advisor, Professor Daniel Scharfstein, for leading my way through biostatistics by introducing me to sophisticated statistical methods. Professor Scharfstein demonstrated to me the crucial qualities of being a successful biostatistician. He encouraged me to become a qualified biostatistician in the future. I am thankful to him for sharing his time with me, giving me research ideas, and correcting my naive mistakes. It is a rare case that a prestigious biostatistician could help a Master's student in this way. Also, I am grateful to my thesis reader, Professor Elizabeth Colantuoni. She was on the committee that granted me admission to the Biostatistics ScM program at JHSPH and has kept close communication with me during the course of my study.

I am thankful to Professors Scott Zeger, Mei-Cheng Wang, Karen Bandeen-Roche, Hongkai Ji, and Ni Zhao for providing me with consistent academic and life support. They helped me with my research projects, encouraged me when I hesitated about my future, and kindly supported all the decisions I made.

I am also thankful to Professor Ellen MacKenzie and Craig Remenapp, as I have learned so much through our collaboration on the MIMIC research project. Through this collaboration, I have learned the importance of biostatistics in addressing an important public health question.

I want to convey my sincere thankfulness to Jingfei Lu for always being considerate and supportive of my research. Her kindness and charisma have influenced me significantly and helped me to become a better person. I am also grateful to Dr. Guanghao Qi, Bonnie Smith, Erjia Cui, Runzhe Li, Zhi Yu, and Jingning Zhang for sharing their invaluable experience as doctoral students and guiding my career path in the future.

Finally, I want to express my gratitude to all the students who participated in the *Essentials of Probability and Statistical Inference* course series. This was my first experience as a teaching assistant. I know that I have made multiple mistakes during weekly labs, but they have never blamed me for them.

Dedication

This thesis is dedicated to my parents, Sa Li, and Lingyun Wang. They led my way to biostatistics and encouraged me during the most desperate moments in my lifetime.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In many randomized trials, patients are scheduled to take assessment at fixed points in time after randomization. In practice, the number and timing of outcome assessments often vary from patient to patient. Importantly, the timing of assessments may be informative in the sense that the distribution of outcomes at the scheduled time for those who are assessed at that time can be different than for those who miss completely or present off-schedule.

To illustrate, consider the *Asthma Research for the Community (ARC)* study. *ARC* was a PCORI-funded pragmatic trial, which enrolled 301 low-income individuals with uncontrolled symptoms of asthma. [1] [2] Each participant was randomized to a control (PT) or intervention (PT + HV) group. The PT group involves usual care plus a web-based training program, which aims to improve the relationship between a patient and their healthcare providers. The PT + HV group adds home visits by community workers. The study was designed to collect asthma control outcomes at 3 months, 6 months, 9 months, and 12 months after randomization. In the PT group, 92%, 72%, 61%, and 92% of the participants completed their 3, 6, 9, and 12 asthma control questionnaires. In the PT + HV group, the corresponding percentages are 92%, 70%, 54%, and 89%. These completion rates do not accurately reflect the variability in the assessment times relative to the targeted assessment times. Figure (1.1) shows how the actual assessment times compared to the planned times for the PT (top panel) and PT + HV (bottom panel) groups. For both groups, we notice that there is a tendency of delayed response for all four assessments. While the distributions of assessment times are similar across treatment groups, this does not imply that the treatment effects, estimated using traditional methods, will be unbiased. This is due to the potentially informative nature of data collection and the fact that the degree of informativeness can differ between groups.

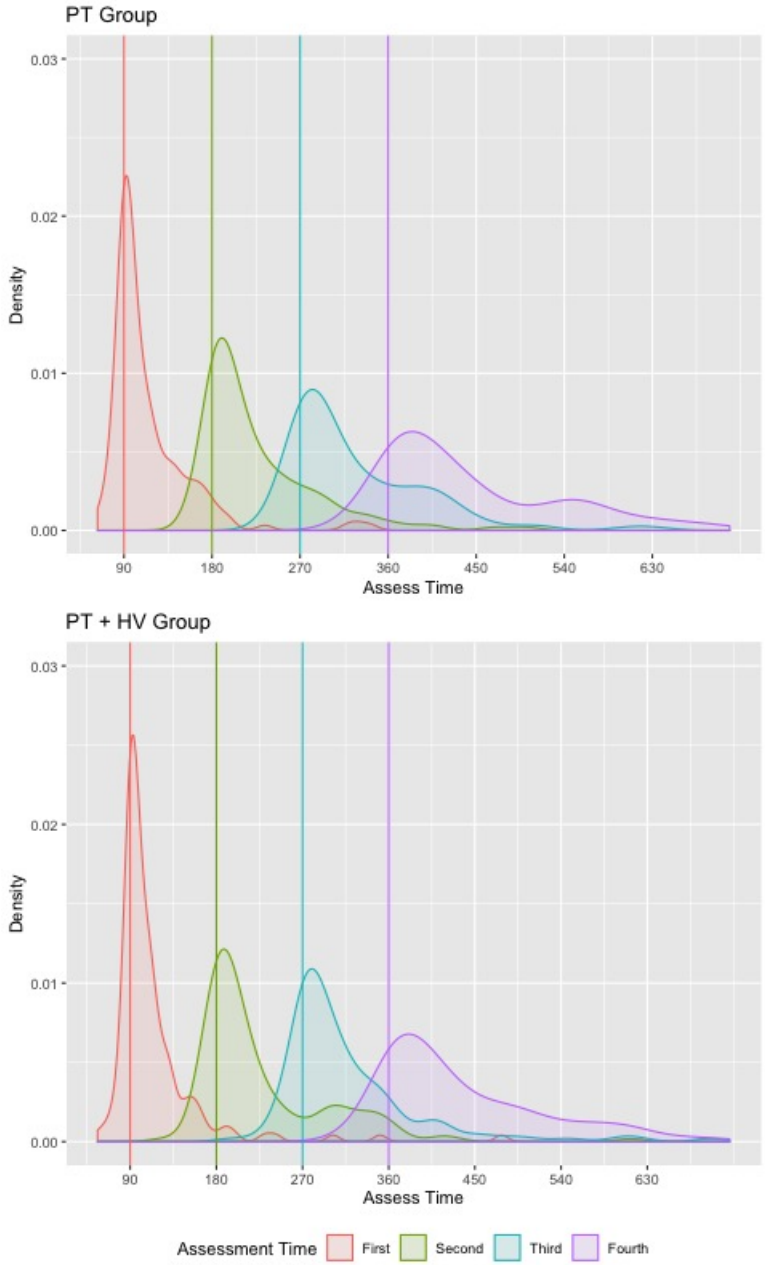

Figure 1.1: The Density of Assessment Times by Visit, Stratified by Treatment Group.

In the presence of irregular assessment times, researchers typically create assessment windows and then apply missing data methods. This approach is unnatural and ad-hoc as the outcomes of patients assessed only one day apart will be treated differently: one within the window (observed) and one outside the window (missing). Rather than dichotomizing in this fashion, a more natural approach is to treat the assessments

as a stochastic process that plays out over time. Towards this end, two broad classes of methods have been proposed. The first class assumes that the assessment-time process is explainable by observable information, i.e., the risk of being assessed at each time t is independent of the outcome at time t given the past history of information collected prior to time t. [8][9][15][16] This assumption will be violated if participants are more likely to be assessed at a given time based on their outcome at that time, even after accounting for their past history. The second class posits a joint model for the outcome and assessment-time processes using shared or dependent random effects. [4] [6] [7] [10] [17] [19] [20] [21] These methods depend on strong distributional assumptions and induce very specific dependence structures between the two processes. Both classes of methods depend on fundamentally untestable assumptions.

In Chapter 2, we introduce the data structure and notation. Chapter 3 discusses our modeling approach. Chapter 4 presents a re-analysis of the *ARC* study. Chapter 5 is devoted to a discussion.

Chapter 2

Data Structure and Notation

We introduce notation for a random individual. Until later, we ignore treatment assignment. For our context, we define random variables that are indexed by time as well as their histories through time. Let $Y(t)$ be the outcome at time t, $N(t)$ be the number of assessments that occur at or prior to time t and $dN(t)$ be the indicator of assessment at time t. Let $\overline{Y^{obs}}(t) = \{Y(s) : dN(s) = 1, 0 \leq s \leq t\}$ and $\overline{N}(t) = \{N(s) : 0 \leq s \leq t\}$ be the history of observed outcomes and observed assessment times through time t, respectively. We use the $t-$ notation below to exclude time t. We denote the history of the observed data for an individual through time t as $\overline{O}(t) = (\overline{N}(t), \overline{Y^{obs}}(t))$. We assume that we observe n independent and identically distributed copies of $\overline{O}(\tau)$, where τ is a fixed time after the last scheduled assessment.

The goal is to use the observed data to draw inference about the mean of $Y(t_1), \ldots, Y(t_K)$, where $t_1 < \ldots, t_K < \tau$ are the protocolized assessment times.

Chapter 3

Modeling Approach

3.1 Assessment Time Processes - Assumptions

Since not all participants are assessed at the protocolized assessment times, assumptions are required. Like the missing data literature, three assumptions have been proposed: assessment completely at random (ACAR), assessment at random (AAR), and assessment not at random (ANAR). [7] Note that the literature uses the "visiting" prefix, but we prefer the "assessment" prefix due to its greater generality. ACAR states that the assessment-time process is completely unrelated to the outcome process. This implies that the distribution of the outcome at time t is the same for those who are at assessed at t and those who are not assessed at t, i.e.

$$dF(Y(t)|dN(t) = 0) = dF(Y(t)|dN(t) = 1) \tag{3.1}$$

where $dF(\cdot|\cdot)$ represents the conditional distribution function. AAR states that assessment at time t is unrelated to the outcome at time t, conditional on the past history observed prior to time t. Alternatively, AAR states that, conditional on the past history observed prior to time t, the distribution of the outcome at time t is the same for those who are not assessed at t and those who are assessed at t. Mathematically, we can write this assumption as:

$$dF(Y(t)|dN(t) = 0, \overline{O}(t-)) = dF(Y(t)|dN(t) = 1, \overline{O}(t-)) \tag{3.2}$$

ANAR states that the outcome at time t is related to assessment at time t even after accounting for the

past history observed prior to time t. That is, the equalities in Equations (3.1) and (3.2) above do not hold. There is an infinite number of ANAR assumptions, making it no way to enumerate.

The AAR and ANAR assumptions are not testable from the observed data. This can be seen by inspecting Equation (3.2). Although the distribution on the right side of the equation is, with large enough samples and sufficient smoothness conditions, estimable from the observed data, the distribution on the left side is not. This is because the outcome at time t is not observed for those who do not show up at that time t. Thus, AAR is an assumption that is equating a fundamentally unknowable distribution to one that is knowable. As a result, there is no way to test AAR. By the same logic, there is no way to test ANAR. In such settings, sensitivity analysis is critical to assess the robustness of study conclusions to unverifiable assumptions.

3.2 Sensitivity Analysis Model

Towards this end, we propose a class of assumptions, indexed by sensitivity analysis parameters, that include AAR and wide collection of ANAR assumptions. We build a class of assumptions using a device called exponential tilting. [3] Specifically, we assume that

$$\begin{aligned} &dF(Y(t)|dN(t)=0,\overline{O}(t-)) \\ &= \frac{dF(Y(t)|dN(t)=1,\overline{O}(t-))\exp\{-q(t,\overline{O}(t-),Y(t);\alpha)\}}{E(\exp\{-q(t,\overline{O}(t-),Y(t);\alpha)\}|\overline{O}(t-))} \end{aligned} \tag{3.3}$$

where $q(t,\overline{O}(t-),Y(t);\alpha)$ is a specified function of its arguments that does not depend on $Y(t)$ only when $\alpha=0$. It is important to notice that when $\alpha=0$, Equation (3.3) reduces to Equation (3.2), i.e., AAR, and when $\alpha\neq 0$, Equation (3.3) encodes a specific ANAR assumption. When the time scale is discrete, it can be shown that (3.3) can be re-expressed as follows:

$$\text{logit}\{P(dN(t)=1|\overline{O}(t-),Y(t))\} = h(\overline{O}(t-);\alpha) + q(t,\overline{O}(t-),Y(t);\alpha), \tag{3.4}$$

where

$$\begin{aligned} h(\overline{O}(t-);\alpha) = &\text{logit}\{P(dN(t)=1|\overline{O}(t-))\}+ \\ &\log\{E(\exp\{-q(t,\overline{O}(t-),Y(t);\alpha)\}|\overline{O}(t-))\} \end{aligned}$$

In our analysis in Chapter 4, we will set

$$q(t, \overline{O}(t-), Y(t); \alpha) = \alpha Y(t) \tag{3.5}$$

Using this sensitivity analysis function, $\exp(\alpha)$ is interpreted as the conditional (on past history) odds ratio of being assessed at time t for individuals who differ by one unit in $Y(t)$. When $\alpha > 0$ (< 0), individuals with higher (lower) levels of $Y(t)$ are more likely to be assessed at time t.

3.2.1 Identification

Under (3.3), $E[Y(t)]$ is identified using the following formula:

$$\begin{aligned}
&E[Y(t)] \\
&= \int_{\overline{o}(t-)} \int_{y(t)} y(t) dF(y(t)|dN(t) = 1, \overline{O}(t-) = \overline{o}(t-)) \times \\
&\qquad \Big\{ P(dN(t) = 1|\overline{O}(t-) = \overline{o}(t-)) + \\
&\qquad \frac{\exp\{-q(t, \overline{o}(t-), y(t); \alpha)\} P(dN(t) = 0|\overline{O}(t-) = \overline{o}(t-)}{E(\exp\{-q(t, \overline{O}(t-), Y(t); \alpha)\}|\overline{O}(t-) = \overline{o}(t-))} \Big\} dF(\overline{o}(t-))
\end{aligned} \tag{3.6}$$

This identification formula suggests that in order to estimate $E[Y(t)]$, we need to estimate the following quantities:

- $dF(y(t)|dN(t) = 1, \overline{O}(t-))$ - observed outcome regression
- $P(dN(t) = 1|\overline{O}(t-))$ - assessment process regression
- $dF(\overline{o}(t-))$

Since $dF(\overline{o}(t-))$ is a marginal distribution, it is natural to estimate to it empirically. We will estimate the observed outcome and assessment process regressions using fully parametric models.

3.2.2 Observed Outcome Regression Model

In the *ARC* study, the primary outcome was asthma control, reflecting symptoms over the week prior to assessment. The outcome is a count variable, taking values $\{0, 1, \ldots, 36\}$. A lower value of the outcome represents better control, with 0 indicating total control and 36 indicating extremely uncontrolled. Figure 2 displays histograms of observed asthma control scores for the PT and PT + HV groups, stratified by planned

assessment time. It is important to note that in each figure there is a spike near zero. This suggests a model that accommodates so-called zero-inflation.

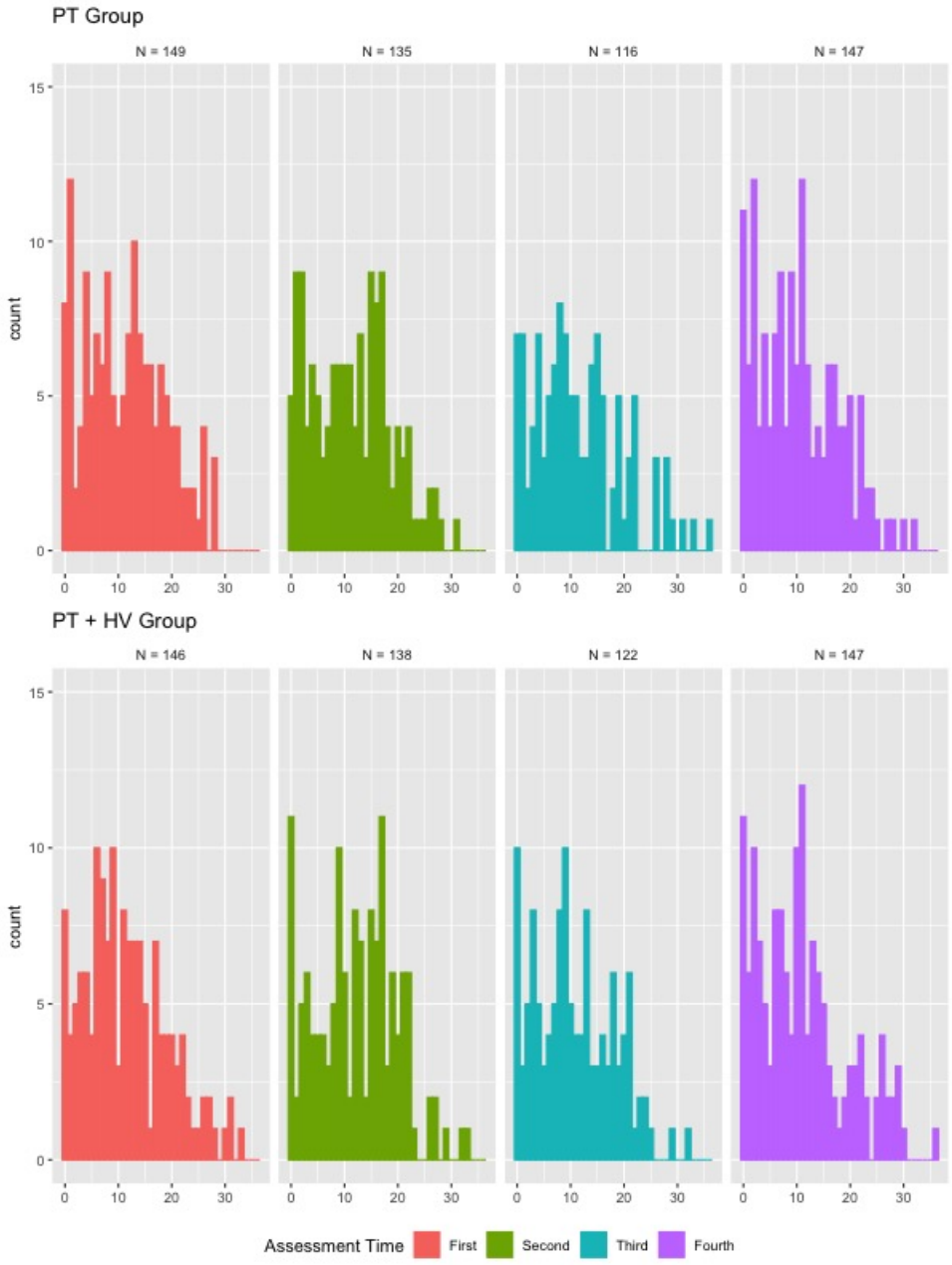

Figure 3.1: Histogram of Observed Asthma Control Scores, Stratified by Planned Assessment Time and Treatment Group.

There are two popular models for handling zero-inflated count data: zero-inflated Poisson (ZIP) and zero-inflated Negative Binomial (ZINB). These models assume that the conditional distribution of $Y(t)$

given $dN(t) = 1$ and $\overline{O}(t-)$ is of the form:

$$\begin{aligned} & dF(y(t)|dN(t) = 1, \overline{O}(t-)) \\ & = \begin{cases} p(t, \overline{O}(t-); \beta) + (1 - p(t, \overline{O}(t-); \beta))h(y(t), t, \overline{O}(t-); \eta) & y(t) = 0 \\ (1 - p(t, \overline{O}(t-); \beta))h(y(t), t, \overline{O}(t-); \eta) & y(t) = 1, 2 \ldots . \end{cases} \end{aligned} \tag{3.7}$$

Both the ZIP and ZINB models assume

$$\text{logit}\{p(t, \overline{O}(t-); \beta)\} = g(t, \overline{O}(t-); \beta) \tag{3.8}$$

where $g(t, \overline{O}(t-); \beta)$ is a specified function of its arguments and β is an unknown vector of coefficients.

The ZIP model assumes

$$h(y(t), t, \overline{O}(t-); \eta) = \frac{\mu(t, \overline{O}(t-); \eta)^{y(t)} \exp\{-\mu(t, \overline{O}(t-); \eta)\}}{y(t)!} \tag{3.9}$$

where $\mu(t, \overline{O}(t-); \eta)$ is a specified non-negative function of its arguments and η is an unknown vector of coefficients. Notice that (3.9) is the Poisson distribution with mean $\mu(t, \overline{O}(t-); \eta)$.

In contrast, the ZINB model assumes

$$\begin{aligned} h(y(t), t, \overline{O}(t-); \eta) = & \frac{\Gamma(y(t) + \theta(t, \overline{O}(t-); \eta_1))}{\Gamma(\theta(t, \overline{O}(t-); \eta_1))\Gamma(y(t) + 1)} \times \\ & \left(\frac{\theta(t, \overline{O}(t-); \eta_1)}{\theta(t, \overline{O}(t-); \eta_1) + \mu(t, \overline{O}(t-); \eta_2)} \right)^{\theta(t, \overline{O}(t-); \eta_1)} \times \\ & \left(\frac{\mu(t, \overline{O}(t-); \eta_2)}{\theta(t, \overline{O}(t-); \eta_1) + \mu(t, \overline{O}(t-); \eta_2)} \right)^{y(t)} \end{aligned} \tag{3.10}$$

where $\Gamma(\cdot)$ is the gamma function, $\theta(t, \overline{O}(t-); \eta_1)$ and $\mu(t, \overline{O}(t-); \eta_2)$ are specified non-negative functions of their arguments and $\eta = (\eta_1, \eta_2)$ is an unknown vector of coefficients.

Notice that (3.10) follows a Negative Binomial distribution with mean

$$\mu(t, \overline{O}(t-); \eta_2) \tag{3.11}$$

and variance

$$\mu(t, \overline{O}(t-); \eta_2)\{1 + \mu(t, \overline{O}(t-); \eta_2)/\theta(t, \overline{O}(t-); \eta_1)\} \tag{3.12}$$

The parameters of the ZIP and ZINB models are typically estimated using maximum likelihood via the EM algorithm. [13]

Figures (3.2) presents Q-Q plots for the Poisson, ZIP and ZINB models for the observed outcomes, stratified by treatment group. The figure shows that the ZINB model provides a superior fit.

3.2.3 Assessment Process Regression Model

To model $P(dN(t) = 1|\overline{O}(t-))$, we will use a model of the form:

$$\text{logit}\{P(dN(t) = 1|\overline{O}(t-))\} = l(t, \overline{O}(t-); \phi) \tag{3.13}$$

where $l(t, \overline{O}(t-); \phi)$ is a specified function of its arguments and ϕ is an unknown parameter vector.

Estimation of ϕ using maximum likelihood can be biased, and such bias can be amplified in the presence of rare events. [11] [12] In particular, predicted probabilities can be overestimated. In the *ARC* study, the number of individuals assessed on any given day tends to be small. Figure 3 shows the daily count of assessments for the PT (top panel) and PT + HV (bottom panel) groups. From the figure, we notice that the maximal number of daily assessments are 17 and 20 in PT and PT + HV groups, respectively. Most of the days have less than five assessments.

One way to correct this bias is to estimate ϕ using Firth's method. [5] Firth's method uses Jeffreys invariant prior to penalizing the logistic regression likelihood function. This penalization helps to correct bias to the first order. [5] Further improvement can be achieved by using the following FLAC procedure proposed by [14]:

1. Estimate ϕ using Firth's method;
2. Compute the hat matrix, which provides a measure of leverage h_i for each observation i;
3. Create a new stacked dataset with three parts:
   (a) Original outcome, original covariates, a new covariate taking on the value 0 for each observation and an individual-specific weight equal to h_i ;

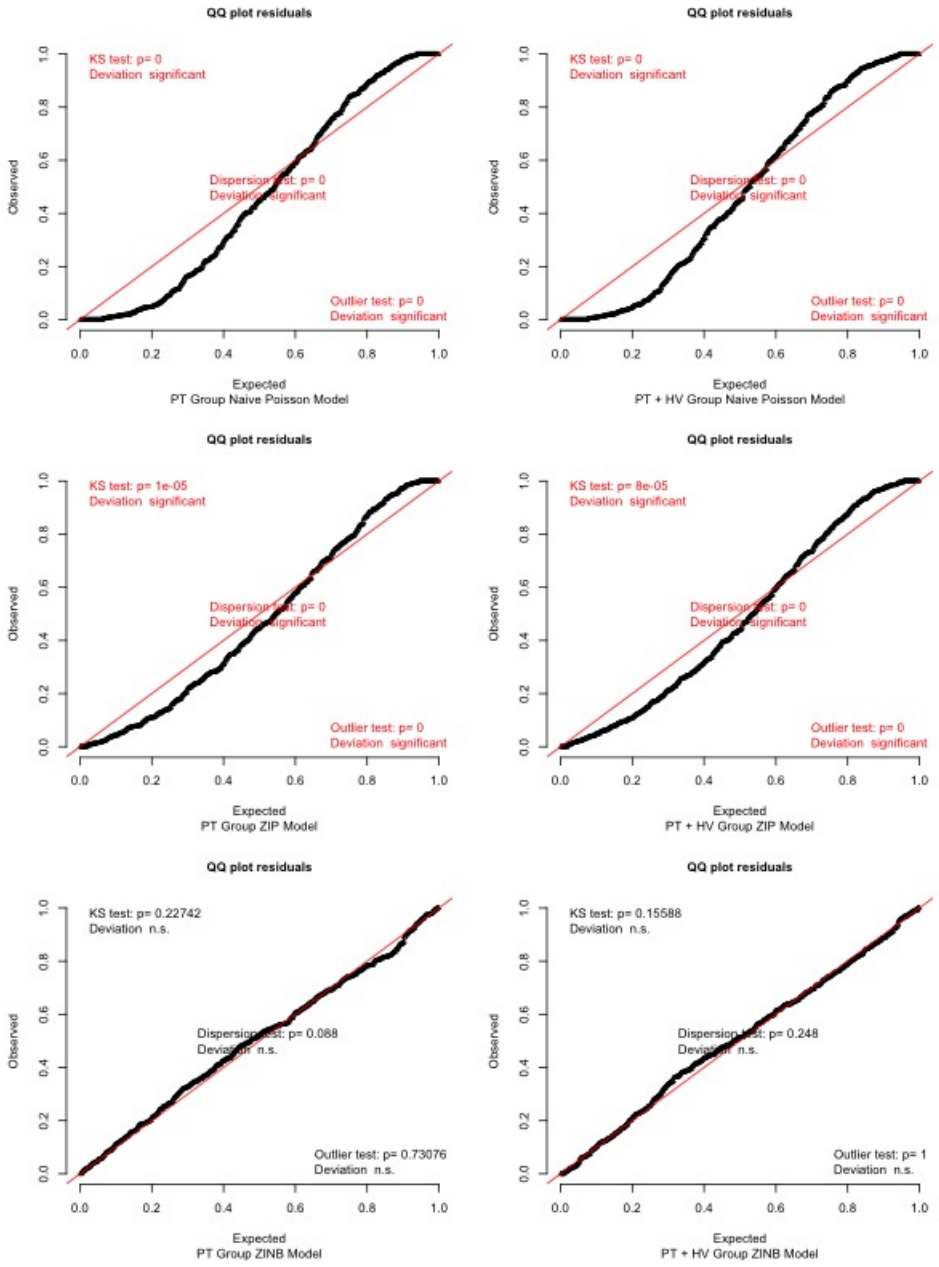

Figure 3.2: Q-Q Plots for Poisson Regression, ZIP and ZINB Models for Observed Outcomes, Stratified by Treatment Group

(b) Original outcome, original covariates, a new covariate taking on the value 1 for each observation and an individual-specific weight equal to $h_i/2$;

(c) Reverse coding of original outcome (1 to 0; 0 to 1), original covariates, a new covariate taking on the value 1 for each observation and an individual-specific weight equal to $h_i/2$;

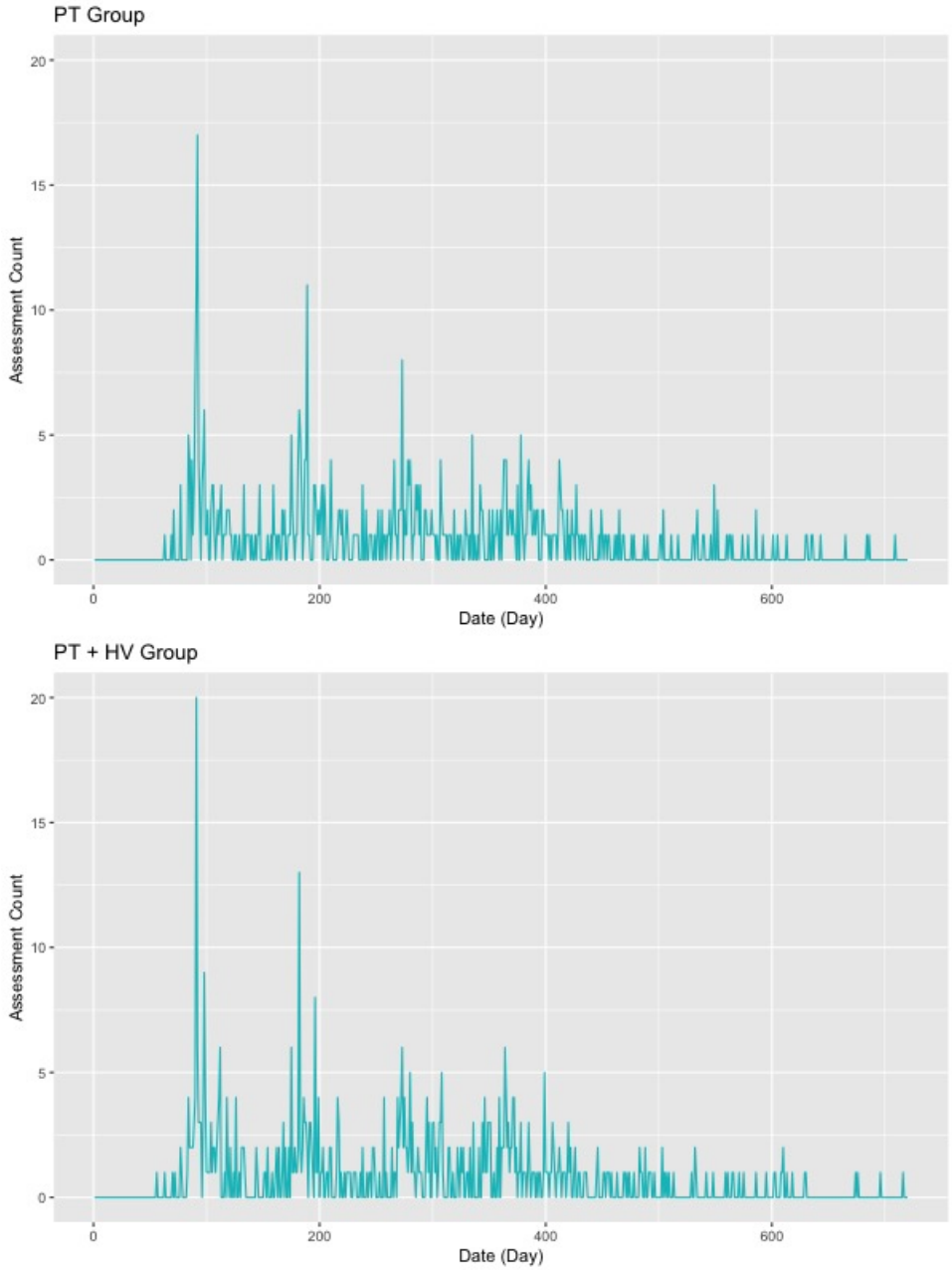

Figure 3.3: Daily Count of Assessments, Stratified by Treatment Group

4. Use the stacked dataset to estimate ϕ using weighted maximum likelihood.

FLAC works by ensuring that the average predicted probability is close to the observed proportion of events. It also moves the log odds ratios closer to 0 as compared to the classic logistic regression model. [14]

3.2.4 Estimation of $E[Y(t)]$

Let $\widehat{\beta}$ and $\widehat{\eta}$ be the estimators of β and η, the parameters of the observed outcome regression model. Let $\widehat{\phi}$ be the estimator of ϕ, the parameter of the assessment process regression model. Using (3.6), estimation of $E[Y(t)]$ proceeds by iterating the following procedure:

1. Draw $\overline{o}(t-)$ from the empirical distribution of $\overline{O}(t-)$;

2. Use $\widehat{\phi}$ to estimate the probability that $dN(t) = 1$ given $\overline{O}(t-) = \overline{o}(t-)$; call the resulting probability $\widehat{P}(dN(t) = 1|\overline{O}(t-) = \overline{o}(t-))$

3. Use $\widehat{\beta}$ and $\widehat{\eta}$ to draw K $y(t)$'s from the estimated distribution of $Y(t)$ given $N(t) = 1$ and $\overline{O}(t-) = \overline{o}(t-)$; denote these draws by $y_k(t)$, $k = 1, \ldots, K$. Estimate $E(\exp\{-q(t, \overline{O}(t-), Y(t); \alpha)\}|dN(t) = 1, \overline{O}(t-) = \overline{o}(t-))$ by

$$\begin{aligned}&\widehat{E}(\exp\{-q(t, \overline{O}(t-), Y(t); \alpha)\}|dN(t) = 1, \overline{O}(t-) = \overline{o}(t-)) \\ &= \frac{1}{K}\sum_{k=1}^{K} \exp\{-q(t, \overline{o}(t-), y_k(t); \alpha)\}\end{aligned}$$

4. Use $\widehat{\beta}$ and $\widehat{\eta}$ to draw one $y(t)$ from the estimated distribution of $Y(t)$ given $N(t) = 1$ and $\overline{O}(t-) = \overline{o}(t-)$; denote this draw by $\widehat{y}(t)$ and compute

$$\begin{aligned}&\widehat{y}(t)\{\widehat{P}(dN(t) = 1|\overline{O}(t-) = \overline{o}(t-)) \\ &+ \frac{\{1 - \widehat{P}(dN(t) = 1|\overline{O}(t-) = \overline{o}(t-))\}\exp\{-q(t, \overline{o}(t-), \widehat{y}(t); \alpha)\}}{\widehat{E}(\exp\{-q(t, \overline{O}(t-), Y(t); \alpha)\}|dN(t) = 1, \overline{O}(t-) = \overline{o}(t-))}\}\end{aligned} \tag{3.14}$$

Steps 1 to 4 are repeated J times and the results from Equation (3.14) are averaged.

The procedure is called the G-computation algorithm. This method was first introduced by James Robins in 1986. [18]

3.2.5 Treatment Comparison

The above models and associated estimation procedures will be employed separately by treatment arm. For combinations of treatment-specific sensitivity analysis parameters, confidence intervals for the difference in asthma control means (PT + HV minus PT) will be computed using non-parametric bootstrap (separate re-sampling of observations within treatment arms).

Chapter 4

Data Analysis

To describe the analysis of the *ARC* study, we introduce the following notation. Let $\kappa(t, \overline{O}(t-)) = \max\{s : dN(s) = 1, 0 \leq s < t\}$ and $Y_{lag}(t) = Y(\kappa(t, \overline{O}(t-))$. In the observed outcome regression model, we let

$$g(t, \overline{O}(t-); \beta) = s(Y_{lag}(t); \beta)$$

$$\log\{\theta(t, \overline{O}(t-); \eta_1)\} = s(t; \eta_{1,1}) + s(t - \kappa(t, \overline{O}(t-)); \eta_{1,2}) + s(Y_{lag}(t); \eta_{1,3})$$

$$\log\{\mu(t, \overline{O}(t-); \eta_2)\} = s(t; \eta_{2,1}) + s(t - \kappa(t, \overline{O}(t-)); \eta_{2,2}) + s(Y_{lag}(t); \eta_{2,3})$$

where $s(\cdot;\cdot)$ is a natural cubic spline function with 3 degrees of freedom, $\eta_1 = (\eta_{1,1}, \eta_{1,2}, \eta_{1,3})$ and $\eta_2 = (\eta_{2,1}, \eta_{2,2}, \eta_{2,3})$. For the assessment regression model, we let

$$l(t, \overline{O}(t-); \phi) = s_1(t; \phi_1) + s_2(t - \kappa(t, \overline{O}(t-)); \phi_2) + s_3(Y_{lag}(t); \phi_3)$$

where $\phi = (\phi_1, \phi_2, \phi_3)$. In our model, $s_1(\cdot;\cdot)$ has 5 degrees of freedom, $s_2(\cdot;\cdot)$ has 10 degrees of freedom, and $s_3(\cdot;\cdot)$ has 3 degrees of freedom. We fit the ZINB model by using the `R` function `zeroinfl` from the `pscl` package. We fit the FLAC method using the `logistf` function in the `logistf` package.

To assess goodness of fit, we used the estimated models to simulate, for each treatment group, a simulated dataset with 2000 individuals. Figure (4.1) displays cumulative distribution functions of simulated and observed asthma control scores for the four assessment times, stratified by treatment group. The figures show a reasonable fit, except possibly for the second assessment in the PT + HV group. Figure (4.2)

displays simulated versus observed densities of first, second, third, and fourth assessment times, stratified by treatment group. The goodness of fit is also reasonable but appears to be better for the PT group than the PT + HV group.

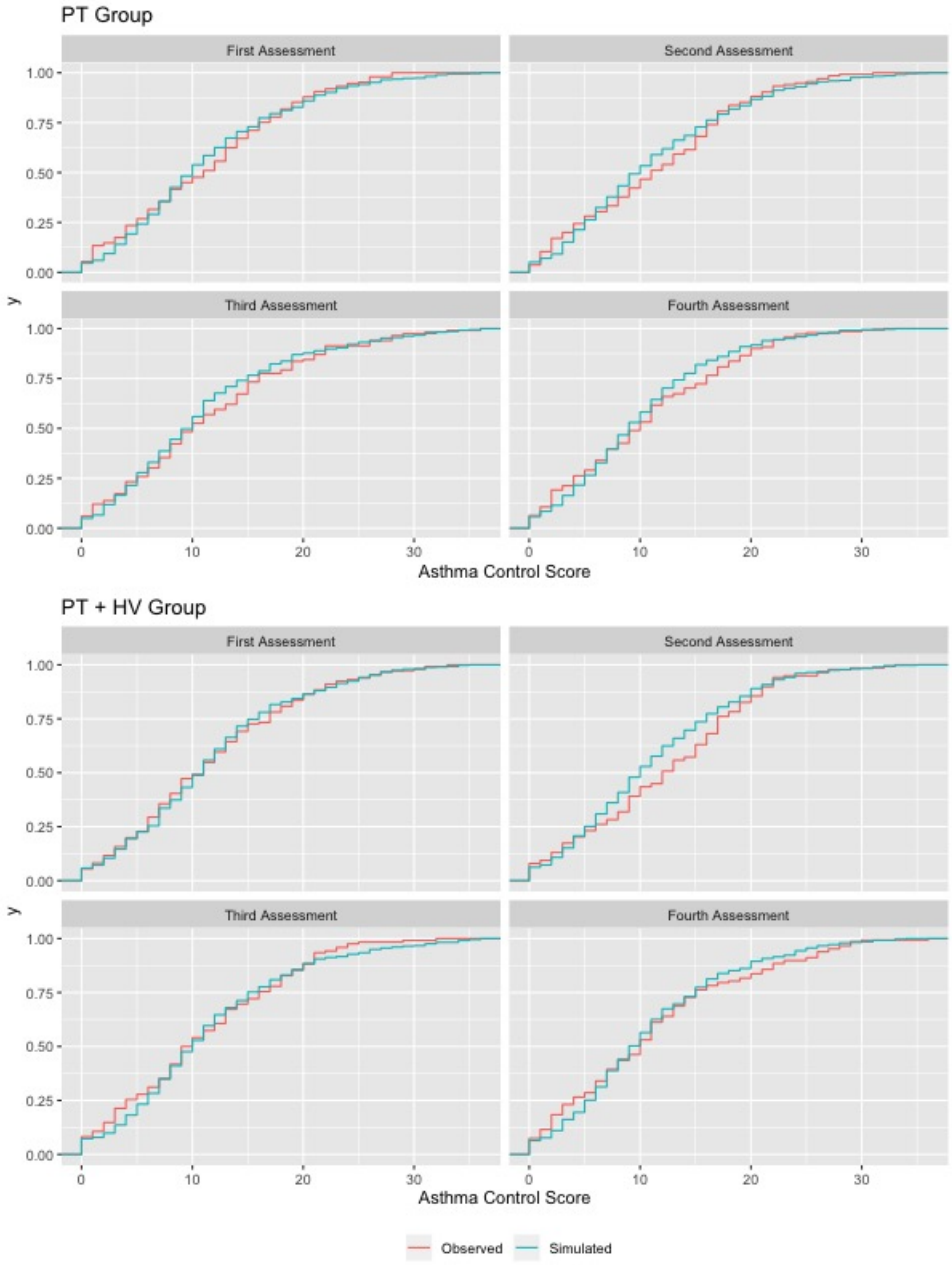

Figure 4.1: Cumulative Distribution Functions for Simulated and Observed Asthma Control Scores by Assessment Time, Stratified by Treatment Group

Figure (4.3) displays treatment-specific estimates of the mean asthma control as a function of time since

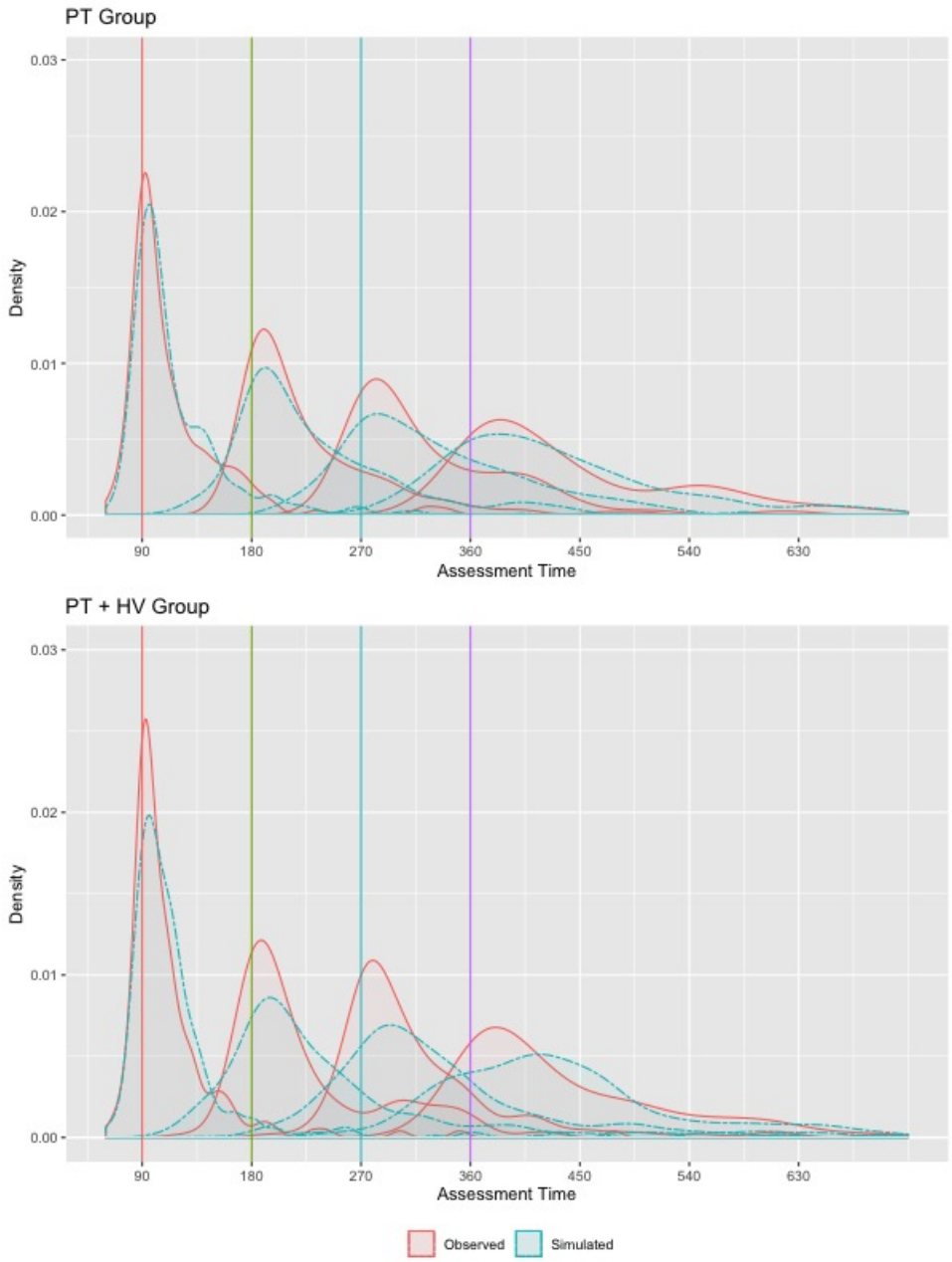

Figure 4.2: Simulated versus Observed Densities of First, Second, Third and Fourth Assessment Times, Stratified by Treatment Group

randomization for various choices of α. Notice that the means decrease with α. This makes sense because positive values of α imply that individuals with worse asthma control are more likely to be assessed at any time t, above and beyond the past history. The figure also shows the observed means for the asthma control scores that were categorized as 3, 6, 9, and 12 months, as in Figure (1.1). These means are similar to those

estimated under AAR.

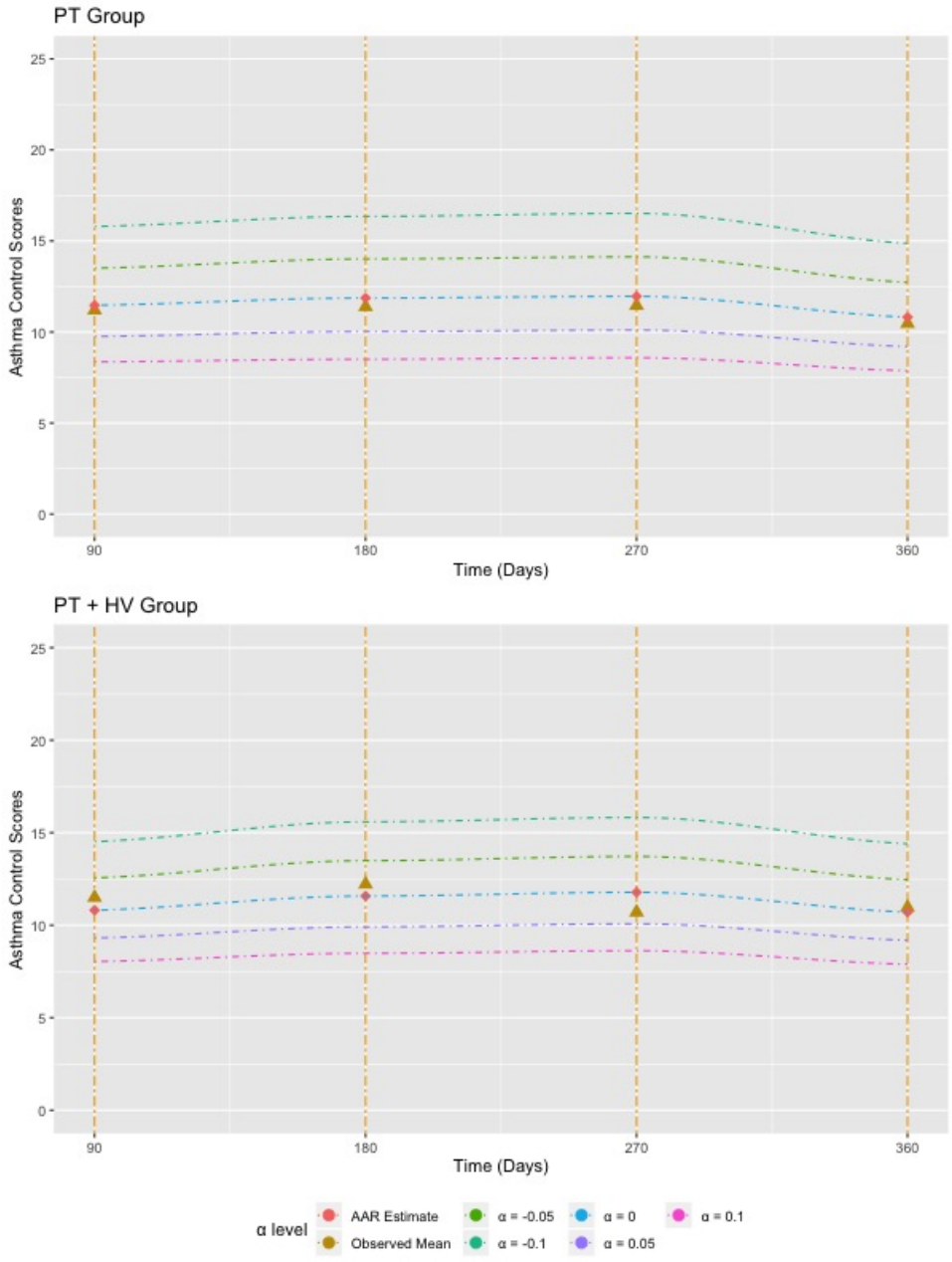

Figure 4.3: Estimated Mean Asthma Control Score by Time for Various Choices of α, Stratified by Treatment Group

Table (4.1) displays the results of the naive analysis that compares treatments with respect to the observed means for the asthma control scores at four assessments. Based on this analysis, there is no evidence of a treatment difference.

Table 4.1: Treatment Comparison Based on Naive Analysis

Assessment	Point Estimate	95% CI
First	0.30	(-1.43, 2.02)
Second	0.83	(-0.97, 2.63)
Third	-0.75	(-2.72, 1.22)
Fourth	0.51	(-1.28, 2.32)

Table 4.2: Treatment Comparisons for Various Combinations of Treatment-Specific Sensitivity Analysis Parameters at 90 Days

		PT + HV				
	α	α = -0.1	α = -0.05	α = 0	α = 0.05	α = 0.1
PT	α = -0.1	-1.30 (-2.52, 0.01)	-3.28 (-4.38,-2.04)	-4.97 (-6.02,-3.84)	-6.43 (-7.45,-5.37)	-7.71 (-8.68,-6.65)
	α = -0.05	1.02 (-0.39, 2.23)	-0.94 (-2.15, 0.06)	-2.67 (-3.79,-1.75)	-4.17 (-5.18,-3.28)	-5.46 (-6.35,-4.56)
	α = 0	3.03 (1.57,4.16)	1.09 (-0.13, 2.10)	-0.65 (-1.74, 0.23)	-2.16 (-3.13,-1.33)	-3.43 (-4.31,-2.66)
	α = 0.05	4.74 (3.39,5.78)	2.79 (1.57,3.70)	1.03 (0.03,1.86)	-0.48 (-1.37, 0.31)	-1.74 (-2.56,-1.05)
	α = 0.1	6.14 (4.79,7.18)	4.22 (3.01,5.01)	2.46 (1.46,3.19)	0.96 (0.08,1.68)	-0.32 (-1.12, 0.32)

Tables (4.2) - (4.5) present treatment comparisons for various combinations of treatment-specific sensitivity analysis parameters, for 90, 180, 270 and 360 days, respectively. As with the naive analysis, analyses conducted assuming AAR holds in both groups (i.e. $\alpha = 0$) provide no evidence of a treatment effect. If, however, there are differential informative assessment mechanisms across the two treatment arms, there is evidence of a difference in mean asthma control between groups; the direction of the difference depends on the difference in signs of the treatment-specific α's. There is no evidence to suggest a differential informative assessment mechanism across two treatment arms. Thus, this analysis does not allow us to conclude a difference between asthma control groups.

Table 4.3: Treatment Comparisons for Various Combinations of Treatment-Specific Sensitivity Analysis Parameters at 180 Days

		PT + HV				
	α	α = -0.1	α = -0.05	α = 0	α = 0.05	α = 0.1
PT	α = -0.1	-0.72 (-2.32, 0.87)	-2.79 (-4.36,-1.36)	-4.70 (-6.18,-3.40)	-6.41 (-7.85,-5.14)	-7.87 (-9.26,-6.62)
	α = -0.05	1.60 (0.16,3.04)	-0.51 (-1.84, 0.81)	-2.44 (-3.69,-1.23)	-4.09 (-5.26,-3.01)	-5.51 (-6.66,-4.53)
	α = 0	3.70 (2.31,5.04)	1.60 (0.41,2.81)	-0.29 (-1.34, 0.77)	-1.97 (-2.94,-1.02)	-3.42 (-4.32,-2.54)
	α = 0.05	5.53 (4.28,6.76)	3.40 (2.32,4.54)	1.52 (0.55,2.49)	-0.12 (-0.97, 0.71)	-1.52 (-2.33,-0.81)
	α = 0.1	7.05 (5.91,8.20)	4.91 (3.95,5.97)	3.02 (2.17,3.93)	1.38 (0.60,2.14)	-0.03 (-0.75, 0.63)

Table 4.4: Treatment Comparisons for Various Combinations of Treatment-Specific Sensitivity Analysis Parameters at 270 Days

		PT + HV				
	α	α = -0.1	α = -0.05	α = 0	α = 0.05	α = 0.1
PT	α = -0.1	-0.65 (-2.66, 1.19)	-2.85 (-4.45,-1.16)	-4.75 (-6.21,-3.26)	-6.46 (-7.85,-5.07)	-7.94 (-9.29,-6.59)
	α = -0.05	1.74 (-0.03, 3.50)	-0.40 (-2.05, 1.23)	-2.32 (-3.74,-0.85)	-4.07 (-5.30,-2.68)	-5.49 (-6.73,-4.21)
	α = 0	3.91 (2.30,5.48)	1.80 (0.21,3.21)	-0.15 (-1.53, 1.19)	-1.84 (-3.05,-0.56)	-3.31 (-4.37,-2.06)
	α = 0.05	5.79 (4.25,7.14)	3.65 (2.17,4.88)	1.73 (0.37,2.85)	0.00 (-1.15, 1.10)	-1.44 (-2.47,-0.39)
	α = 0.1	7.30 (5.83,8.50)	5.22 (3.75,6.24)	3.27 (1.94,4.21)	1.54 (0.41,2.46)	0.08 (-0.91, 0.97)

Table 4.5: Treatment Comparisons for Various Combinations of Treatment-Specific Sensitivity Analysis Parameters at 360 Days

		PT + HV				
	α	α = -0.1	α = -0.05	α = 0	α = 0.05	α = 0.1
PT	α = -0.1	-0.46 (-1.89, 1.29)	-2.41 (-3.82,-0.75)	-4.23 (-5.67,-2.64)	-5.75 (-7.24,-4.28)	-7.03 (-8.46,-5.64)
	α = -0.05	1.69 (0.38,3.22)	-0.27 (-1.46, 1.13)	-2.04 (-3.19,-0.72)	-3.57 (-4.75,-2.35)	-4.88 (-6.05,-3.70)
	α = 0	3.58 (2.33,4.93)	1.65 (0.49,2.89)	-0.11 (-1.16, 1.05)	-1.64 (-2.59,-0.58)	-2.94 (-3.89,-1.93)
	α = 0.05	5.20 (3.98,6.38)	3.26 (2.13,4.37)	1.50 (0.49,2.51)	-0.02 (-0.93, 0.89)	-1.30 (-2.13,-0.41)
	α = 0.1	6.53 (5.33,7.64)	4.55 (3.48,5.58)	2.83 (1.84,3.71)	1.29 (0.42,2.13)	0.02 (-0.78, 0.84)

Chapter 5

Discussion

In this thesis, we introduced a sensitivity analysis methodology for evaluating the effect of treatment in randomized trials in the presence of potentially informative assessment times. We developed this methodology within the context of the *Asthma Research for the Community (ARC)* study. In applying the methodology, we addressed two key challenges: zero-inflated outcomes and rare events.

The methodology relies on the specific of treatment-specific sensitivity analysis parameter. This parameter is interpreted as a regression coefficient in a logistic regression model. Ideally, the range of plausible values of these parameters needs to be specified in conjunction with subject matter experts. This can be a difficult exercise and is best executed prior to data analysis.

A key limitation of our approach is its reliance on fully parametric models. It would be useful to develop a semiparametric extension of our methodology. It would also be useful to extend the methodology to allow the inclusion of auxiliary covariates.

Bibliography

[1] Apter, A. J., Bryant-Stephens, T., Morales, K. H., Wan, F., Hardy, S., Reed-Wells, S., ... Localio, R. Using IT to improve access, communication, and asthma in African American and Hispanic/Latino adults: Rationale, design, and methods of a randomized controlled trial. *Contemporary Clinical Trials*, 44:891–921, 2015.

[2] Apter, A. J., Localio, A. R., Morales, K. H., Han, X., Perez, L., Mullen, A. N., ... Bryant-Stephens, T. Home visits for uncontrolled asthma among low-income adults with patient portal access. *The Journal of Allergy and Clinical Immunology*, 144(3):846–853, 2019.

[3] Barndorff-Nielsen, O. E. and Cox, D. R. *Inference and Asymptotics*. Chapman & Hall, 1994.

[4] Chen, Y., Ning, J., Cai, C. Regression analysis of longitudinal data with irregular and informative observation times. *Biostatistics*, 16(4):727–739, 2015.

[5] Firth, D. Bias Reduction of Maximum Likelihood Estimates. *Biometrika*, 80(1):27–38, 1993.

[6] Huang, C. Y., Wang, M. C., Zhang, Y. Analysing panel count data with informative observation times. *Biometrica*, 93(4):763–775, 2006.

[7] Liang, Y., Lu, W., Ying, Z. Joint modeling and analysis of longitudinal data with informative observation times. *Biometrics*, 65(2):377–384, 2009.

[8] Lin, D. Y. and Ying, Z. Semiparametric and nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association*, 96(453):103–113, 2001.

[9] Lin, H., Scharfstein, D. O., Rosenheck, R. Analysis of longitudinal data with irregular outcome-dependent follow-up. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):791–813, 2004.

[10] Liu, L., Huang, X., O'Quigley, J. Analysis of longitudinal data in the presence of informative observational times and a dependent terminal event, with application to medical cost data. *Biometrics*, 64(3):950–958, 2008.

[11] McCullage, P. and Nelder, J. A. *Generalized Linear Models*. 2nd ed. New York: Chapman & Hall, 1989.

[12] McCullagh, P. *Tensor Methods in Statistics*. New York: Chapman & Hall, 1987.

[13] Minami, M., Lennert-Cody, C. E., Gao, W., Román-Verdesoto, M. Modeling shark bycatch: The zero-inflated negative binomial regression model with smoothing. *Fisheries Research*, 84(2):210–221, 2007.

[14] Puhr, R., Heinze, G., Nold, M., Lusa, L., Geroldinger, A. Firth's logistic regression with rare events: accurate effect estimates and predictions? *Statistics in Medicine*, 36:2302–2317, 2017.

[15] Pullenayegum, E. Multiple outputation for the analysis of longitudinal data subject to irregular observation. *Statistics in Medicine*, 35(11):1800–1818, 2016.

[16] Pullenayegum, E. and Feldman, B. Doubly robust estimation, optimally truncated inverse-intensity weighting and increment-based methods for the analysis of irregularly observed longitudinal data. *Statistics in Medicine*, 32(6):1054–1072, 2013.

[17] Pullenayegum, E. and Lim, L. Longitudinal data subject to irregular observation: A review of methods with a focus on visit processes, assumptions, and study design. *Statistical Methods in Medical Research*, 25(6):2992–3014, 2016.

[18] Robins J. A new approach to causal inference in mortality studies with a sustained exposure period - aplication to control of the healthy worker survivor effect. *Math Model*, 7(9-12):1393–1512, 1986.

[19] Ryu, D., Sinha, D., Mallick, B., Lipsitz, S., Lipshultz, S. Longitudinal studies with outcome-dependent follow-up. *Journal of the American Statistical Association*, 102(479):952–961, 2007.

[20] Sun, J., Sun, L., Liu, D. Regression analysis of longitudinal data in the presence of informative observation and censoring times. *Journal of the American Statistical Association*, 102(480):1397–1406, 2007.

[21] Sun, L., Song, X., Zhou, J., Liu, L. Joint analysis of longitudinal data with informative observation times and a dependent terminal event. *Journal of the American Statistical Association*, 107(498):688–700, 2012.