The Impact of Inherent Measurement Error in the Hemagglutination Inhibition Assay for the Evaluation of Vaccine Immunogenicity

By

Aaron Xian Ti Lee

Thesis
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfilment of the requirements
for the degree of

MASTER OF SCIENCE
in

Biostatistics
August 31, 2021
Nashville, Tennessee

Approved:
Andrew J. Spieker, Ph.D.
Robert A. Greevy, Ph.D.

Copyright © 2021 by Aaron Xian Ti Lee
All Rights Reserved

To the Lord God Almighty, whom I owe my deepest gratitude for this achievement

and

To my parents, who never fail to keep me on track and whose support throughout this program I will always remember

ACKNOWLEDGEMENTS

This work would not have been possible without the incredible help, mentorship, and advice of my thesis instructor, Dr. Andrew Spieker, since the beginning of this research. I owe to him an immense level of gratitude for his insightful ideas, organizational wisdom, and flexibility in accommodating all of my doubts, emergencies, and growing enthusiasm for this project.

In addition, I would like to thank Dr. Robert Greevy for his willingness to be my second reader, as well as for being a highly supportive and encouraging faculty member throughout my educational experience in the Biostatistics Graduate Program. Not forgetting, my sincerest thanks to all of my professors for their knowledge, wisdom, and assistance in their classes as well as to my graduate cohort – Caroline Birdrow, Jamie Joseph, Justin Jacobs, Kaidi Kang, Shengxin Tu, Tianyi Sun, and Xiangyu Ji – for their immense help and friendships throughout our studies these past two years.

Special thanks are due to Dr. Natasha Halasa for her kindness and generosity in granting me access to the data set for the Phase I clinical trial of the trivalent inactivated influenza vaccine among pediatric solid organ transplant patients. I would not have been able to apply the knowledge gained from my research without this wonderful opportunity you have given to me.

I would also like to give a shoutout to University Catholic: to my chaplain, Fr. Gervan Menezes, my campus minister, Nick Aguirre, and every single one in this blessed community for their endless spiritual support, laughter, great friendships, and encouragement to truly pursue excellence in all areas of my life. Your presence in my life since arriving at Vanderbilt is truly inspirational and life-giving!

Finally, but not least of all, I would like to thank my parents not only for their immense support throughout my time at Vanderbilt and being away from home these past two years, but for their never-ending words of encouragement to stoke the drive needed for me to persevere throughout conducting my research and writing this thesis. I hope to never forget the sacrifices you have made for me to achieve this monumental occasion in my educational journey, and both the tough love as well as affection that you shower upon me every day of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviations

HAI	Hemagglutination inhibition assay
RBCs	Red blood cells
SDF	Serial dilution factor
ULQ	Upper limit of quantification
LLQ	Lower limit of quantification
v/v	Volume per volume, or volume of liquid solute per 100 mL of solution
GMT	Geometric mean titer
GMTR	Geometric mean titer ratio
CS	Control group skewness
TS	Treatment group skewness
SS	Sample size per treatment group
TTP	True treatment parameter
MTE	Mean log-transformed treatment estimate
ESD	Empirical standard deviation of treatment estimate
ASE	Average standard error of treatment estimate
CP	Coverage probability of treatment estimate
P	Power of simulation study
RE	Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**
LCO	Log-cumulative odds of treatment parameter
CO	Cumulative odds of treatment parameter
COR	Cumulative odds ratio of treatment parameter
SOT	Solid organ transplant
SD	Standard-dose
HD	High-dose
TIV	Trivalent inactivated influenza vaccine
CI	Confidence interval

Symbols

n	Total sample size
K	Total number of observed titer categories
λ	Shape parameter of the Weibull data generating distribution
σ_{CP}	Precision, or standard error of the CP
α	Significance level, or Type I error rate
n_S	Number of simulations
e	Exponentiation
m	Limiting sample size
n_i	Sample size of the i^{th} ordinal category

Simple linear regression

β_1	True log-transformed treatment effect
$\hat{\beta}_1$	Log-transformed treatment estimate
$\overline{\hat{\beta}}_1$	Mean log-transformed treatment estimate
$\sigma_e\left(\hat{\beta}_1\right)$	Estimated standard deviation of the treatment estimate
$\overline{\sigma}\left(\hat{\beta}_1\right)$	Average standard error of the log-transformed treatment estimate (simple linear)

Proportional odds regression

β_1	True log-cumulative odds of the treatment parameter
$\hat{\beta}_1$	Estimated log-cumulative odds of the treatment parameter
$\overline{\hat{\beta}}_1$	Mean log-cumulative odds estimate of the treatment parameter
$\sigma_e\left(\hat{\beta}_1\right)$	Estimated standard deviation of the log-cumulative odds of the treatment parameter
$\overline{\sigma}\left(\hat{\beta}_1\right)$	Average standard error of the log-cumulative odds of the treatment parameter

CHAPTER 1

INTRODUCTION

1.1 Immunology and Vaccinology

Vaccines are a major development in modern medicine that originated from early immunization methods to combat infectious, deadly diseases. The first method of inducing viral immunization was a form of inoculation known as variolation, where viral material from a patient infected with a disease is introduced into a person who does not yet suffer from the disease [8]. The person would still suffer the ill effects of the disease, but to a lesser degree than the patient who acquired the infection naturally. An immunized individual through this method would be at least partially protected from future symptomatic disease after recovery. Variolation in the general sense is also known as smallpox inoculation, as it is most associated with smallpox and its eventual worldwide eradication from the population. This in part was due to Edward Jenner's pioneering work in 1798 that demonstrated the achievement of immunological protection from smallpox exposure either through natural cowpox infection or variolation with cowpox viral material from an infected patient [8].

Variolation was not without risks due to possible complications that can arise from the person-to-person inoculation of the virus; between the late 19th and early 20th centuries, safer methods of vaccination were being studied in an experimental setting. Specifically, the first prototypes of the two broad vaccine groups, i.e., live attenuated and inactivated, were developed. About 80 years after Jenner's publication of his work on variolation, Pasteur's observation of the lack of fowl cholera cases among chickens inoculated with cultures left out for a prolonged period led him to experiment with attenuating bacteria by exposing them to adverse conditions to weaken the pathogenic potency of the bacteria [4]. His work led to the first live attenuated vaccine that targeted rabies [4]. Approximately a decade later, Daniel Elmer Salmon and his research assistant Theobald Smith pioneered and published work on immunization of pigeons through inoculation with inactivated, or killed, *Salmonella enterica* bacteria from their studies on hog cholera [4]. Though erroneously attributing *Salmonella enterica* to the cause of hog cholera, their discovery of the efficacy of inactivated bacteria in inducing immunization led to the future development of inactivated vaccines against typhoid, the plague, and cholera [4].

Many inventions of modern-day vaccines could be traced back to these initial developments that arose from physical observations of the presence or absence of the effects of viruses and bacteria on those who were infected. However, with this same methodology, Pulendran and Ahmed noted that many successful vaccines were empirical in their development; deep knowledge and understanding of the underlying immunological mechanisms that help to elicit protective immunity in an organism was barely present or even non-existent in early research and invention of vaccines [12]. The difference in the available technologies to properly study these mechanisms between the present day and the late 18th to the early 20th centuries easily explains this perspective of vaccine development. Nevertheless, due to this anomaly, immunology and vaccinology evolved in vastly different directions though these two disciplines share a common origin from these early developments [12]. However, recent advancements in the understanding of the human innate immune system, its relationship to the introduction of a vaccine into the body, and its role in inducing adaptive immunity, i.e., immunity against a disease not acquired by natural exposure, have begun to innovate current approaches to vaccine development [12]. Incorporation of the knowledge of vaccine-induced immunological mechanisms from the study of the innate immune system into the design of new vaccines is but one of many changes to the discipline.

Public awareness of infectious diseases as well as a rising need to maintain public health in a population inevitably arose with the growing availability of vaccines as well as the growing prevalence of mutation of strains of existing viruses. With increasing concerns to ensure the safety, usability, and efficacy of vaccines before its license for distribution and human use, extensive guidelines for the rigorous evaluation of vaccines throughout their development were outlined by governing bodies such as the World Health Organization (WHO), the European Medicines Agency (EMA), and the United States Food and Drug Administration (USFDA) [14]. These guidelines are especially applicable to novel vaccines, defined as “either the first of its kind based on the mechanism of protection or as the first vaccine for a disease.” [14] As a vaccine candidate becomes approved for testing in a clinical trial, it would sequentially undergo through the standard clinical trial phases of Phase I, Phase II, and Phase III depending on its success in achieving certain thresholds in several criteria, among which include toxicity data and immunogenicity response [14]. Phase IV of a vaccine clinical trial would follow a successful licensure of the vaccine for public use after Phase III, where the safety and efficacy of the vaccine among the population as it is being administered is monitored continuously [14].

1.2 Hemagglutination Inhibition Assay

Immunogenicity assays are analytical procedures to assess the amount of an antibody or antigen that would either suppress or induce respectively an immune response in a person, and are typically conducted using patient blood serum samples. They are often part of the initial evaluation of biologic activity of a vaccine candidate in clinical trials, specifically in Phase I and Phase II as this is when the vaccine candidate is first administered to human subjects [14]. This phase is important not only in determining the safety levels of the vaccine candidate in human use, but also in evaluating its immunological effectiveness through the collection of immune response data as aforementioned in clinical trials criteria.

A popular immunogenicity assay that is often utilized in evaluating vaccine immunogenicity is the hemagglutination inhibition assay (HAI), often utilized in influenza vaccine clinical trials [9, 15]. The HAI has several distinct advantages over other available assays that has led to its extensive use in the analysis of influenza vaccine immunogenicity [9, 11]:

- The HAI is inexpensive and simple to carry out.
- It can be used to test for specific influenza strains.
- Results obtained from the HAI are highly reliable.

The HAI leverages the capacity of hemagglutinin (a glycoprotein found on the surface of flu viruses) to bind to the sialic acid receptors on the membranes of red blood cells (RBCs), causing the RBCs to clump together in a process known as agglutination [9]. This process can be interrupted with the presence of agglutination-inhibiting substances such as influenza antibodies that bind to the virus antigenic sites, thus preventing the virus-RBC binding from occurring [3, 9]. Depending on the presence of, as well as the concentration of antibodies and viruses in blood serum relative to each other, the visual effects of RBC agglutination or otherwise can be physically observed and analyzed.

The HAI is typically carried out with a microtiter plate as shown in Figure 1, with U-bottom or V-bottom wells being the most common well types used in the HAI. Each row of wells is normally used for observations either from the blood sera of different patients or from different virus strains to be examined [5, 9, 11]. In the case of examining blood sera from different patients, the following illustrate the basic outline of performing the HAI [5, 9, 11, 15]. A starting serum dilution titer, a serial dilution factor (SDF), and an upper limit of quantification (ULQ) based on the SDF are initially decided. For

each patient, serum is obtained from the blood circulatory system and several sera dilutions are prepared from the extracted serum. This is done by first adding a small, fixed amount of serum into each well across a specific row. Diluent is added to the first well and mixed with the serum to achieve the selected starting dilution titer, e.g., to make a starting dilution titer of 1:10, 1 part serum is mixed with 9 parts diluent.

Sera in subsequent wells are then serially diluted in multiplicative increments based on the chosen SDF up to the ULQ, e.g., using the previous starting dilution titer of 1:10 and a two-fold SDF, the next titers across the chosen row would then be 1:20, 1:40, 1:80, etc. A fixed amount of the influenza virus sample is then added to each well and is mixed thoroughly with the dilution. Once all serial dilution titers are prepared for each patient, the plate is covered and allowed to incubate at room temperature for at least 30 minutes.

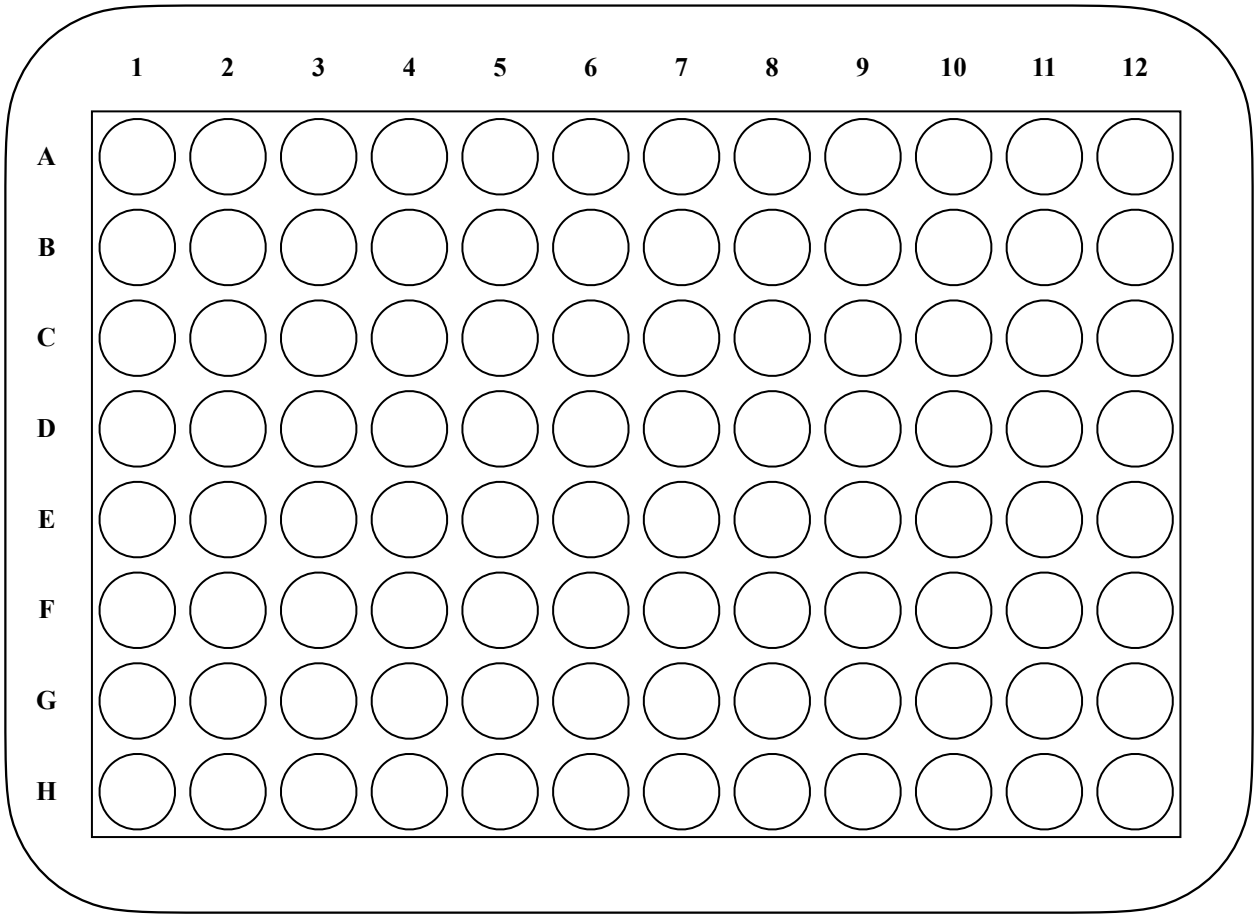

Figure 1: A 96-well microtiter plate in an 8 × 12 array. Among microtiter plates in various sizes, this array is most commonly used [5, 9, 11, 15]. Rows are usually for distinguishing different serum samples while columns indicate the dilution level based on the SDF.

After incubation, RBCs are added into each well and mixed with the dilution; the RBCs can come from mammals, fowl, or humans. The source of RBCs used influences the HAI components necessary to perform it [9, 11, 15]. Kaufmann et al. summarizes the various HAI preparations required and the appearance in the case of non-agglutination based on the RBC type in Table 1 below [5]:

HAI characteristics	RBC types	
	Fowl	**Mammals**
Selected species	Chicken, turkey	Guinea pig, horse, human
RBC concentration (v/v)	0.75%	1%
Microtiter plate shape	V-bottom	U-bottom
Second incubation duration	30 minutes	60 minutes
Non-agglutination appearance	Button	Halo

Table 1: Summary of HAI preparation for different RBC types

The plate is covered when the RBCs are well mixed, and it is allowed to incubate at room temperature for the second time at the duration specified in Table 1 based on the RBC type. After the second incubation period, the wells can be examined for the presence of agglutination. Using the example of fowl RBCs and V-shaped wells according to the standard HAI practices, observations of individual wells in the microtiter plate would demonstrate either the presence, partial, or absence of agglutination for each well based on the following characteristics [5, 11]:

- A well with no agglutination would display a red button in the middle when it is upright, indicating the concentration of RBCs at the bottom of the well. Tilting the well at an angle between 60° and 90° would cause the RBCs to run or streak down the side of the well and pool at the bottom edge.
- A well with partial agglutination would show a pale reddish diffusion with a smaller red button in the middle when it is upright; a smaller run or streak down the side of the well without pooling at the bottom edge would be observed when the well is tilted.
- A well with complete agglutination would exhibit a concentrated reddish diffusion either with or without a tiny red button in the middle, and no running or streaking down the side of the well when it is tilted.

Wells that indicate either partial or complete agglutination would be considered as having agglutination present in the analysis. A visual summary of HAI well observations as previously described is shown in Figure 1 below. Wells are examined both in an upright and tilted position in order to detect the level of agglutination if present.

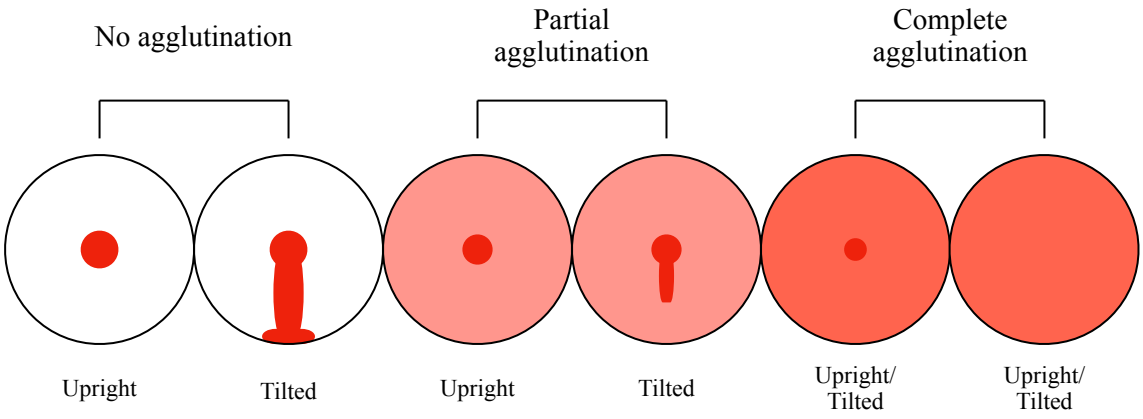

Figure 2: Possible visual displays of HAI wells using fowl RBCs, with well positions

Each patient's serum HAI titer can be determined through the examination of wells; for each row corresponding to the dilution samples from each patient, the patient's serum HAI titer would be the reciprocal of the highest dilution titer that exhibits complete absence of agglutination [5, 9, 15]. The following highlight several examples of identifying patient serum HAI titer under different observational circumstances [5, 9, 15]:

- A patient's serum dilution sample shows no agglutination up to a 1:320 dilution, with the 1:640 dilution sample displaying partial agglutination. The patient's serum HAI titer would thus be 320, the reciprocal of 1:320.
- If all of a patient's serum dilution samples exhibit no agglutination, the patient's serum HAI titer is set to the ULQ of the HAI, e.g., a patient's serum HAI titer is 1280 if sample shows no agglutination up to the ULQ of 1:1280.
- If agglutination is present in a patient's first serum dilution sample, the patient's serum HAI titer would often be the next lower dilution titer before the starting dilution titer, e.g., with a starting titer of 1:10 and an SDF of 2, a patient's serum HAI titer would be 5 if agglutination is present in the 1:10 dilution.

Figure 3 illustrates a visual representation of determining the serum HAI titer of each patient based on the possible settings as previously described, under a starting dilution titer of 1:10 and an SDF of 2. For simplicity, agglutinated wells show complete agglutination and all the wells are upright.

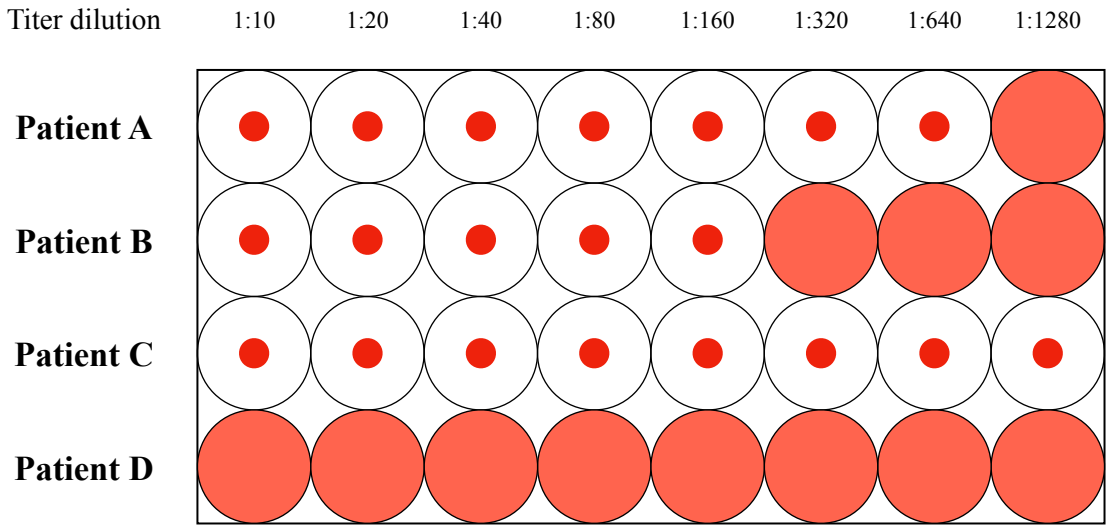

Figure 3: Visual analysis of a cross-section of a sample HAI from four different patients. The serum antibody titer of each patient from the above observations under the established guideline would be A = 640, B = 160, C = 1280, D = 5.

Serum extraction and the HAI are conducted several times throughout a clinical trial, once before the administration of the vaccine to the selected treatment group, and at least once after a waiting period following the administration of the vaccine depending on the clinical objective. Statistical measures and analyses can be performed with the patient serum HAI titer data at each collection point throughout the trial. Each statistical measure can be analyzed with various approaches, including parametric tests, non-parametric tests, and regression modeling. The measures being studied as well as the statistical approaches utilized to do so are dependent on the clinical objective and available resources. The conventional measures that are studied include, but are not limited to:

- The geometric mean titer (GMT) of treatment groups and/or geometric mean titer ratio (GMTR) between treatment groups [1, 9, 11, 15]
- Seroprotection, defined as having a post-vaccination serum HAI titer meeting or exceeding a certain threshold, e.g., ≥ 1:40 as typically used in influenza vaccine studies [1, 6]
- Seroconversion, defined as having a post-vaccination titer that is at least a certain fold increase from baseline e.g., at least a four-fold increase as typically used in influenza vaccine studies [1, 9, 15]

1.3 **Assay Limitations**

Despite the usefulness and convenience of the HAI in analyses of vaccine immunogenicity, it is not without disadvantages from a statistical perspective. Perhaps the most notable limitation of the HAI is the discretization of serum titers through its dilution process. The SDF pre-selected before the HAI is carried out is always a natural number as titer calculations are made convenient and the dilution samples are easy to prepare; the most frequent choice of SDF is 2, though it is not unheard of for research labs to use an SDF of 3 or even 4 [10].

This discretization method categorizes a patient's latent true serum titer by rounding it down to the closest observable dilution titer available in the HAI [15]. Using the previous example in Page 6 of a patient's serum dilution sample showing no agglutination up to a 1:320 dilution with the 1:640 dilution sample displaying partial agglutination, while the observed serum HAI titer would be 320, the patient's actual latent serum HAI titer in this example would lie anywhere in the range of [320, 640), even when under the assumption that no other sources of measurement error are present. Thus, the observed serum HAI titer obtained by this method will almost always be lower than the true (and latent) titer, assumed to be continuous. As the HAI cannot directly measure the patient's true titer due to this discretization, there is inherent measurement error incorporated into the assay that leads to loss of information.

Furthermore, the degree of bias and information loss is clearly linked to the magnitude of the pre-specified SDF. The increasing degree of information loss as the SDF increases can be observed with a simple simulation as follows: data for latent titers were simulated for both a control (untreated) group and a vaccine (treated) group, with 250 patients in each group. The data were generated using a Weibull distribution, with the vaccine group having more patients with higher titers overall assuming that there is a positive treatment effect. The data of both groups were combined, and the observed titers for each patient were obtained using SDFs of 2, 3, and 4, as well as a minimum dilution titer of 5. The distribution of observed titers as well as the observed GMT of the combined groups for each SDF were then evaluated.

The results of the simple simulation of titer discretization are shown in Figure 4. Assuming a specific distribution of theoretical latent true titers, the distribution of the observed titers becomes less reflective of the former and the GMT becomes increasingly biased as the SDF increases. The greater information loss with each higher SDF would lead to a higher likelihood of specious data analyses.

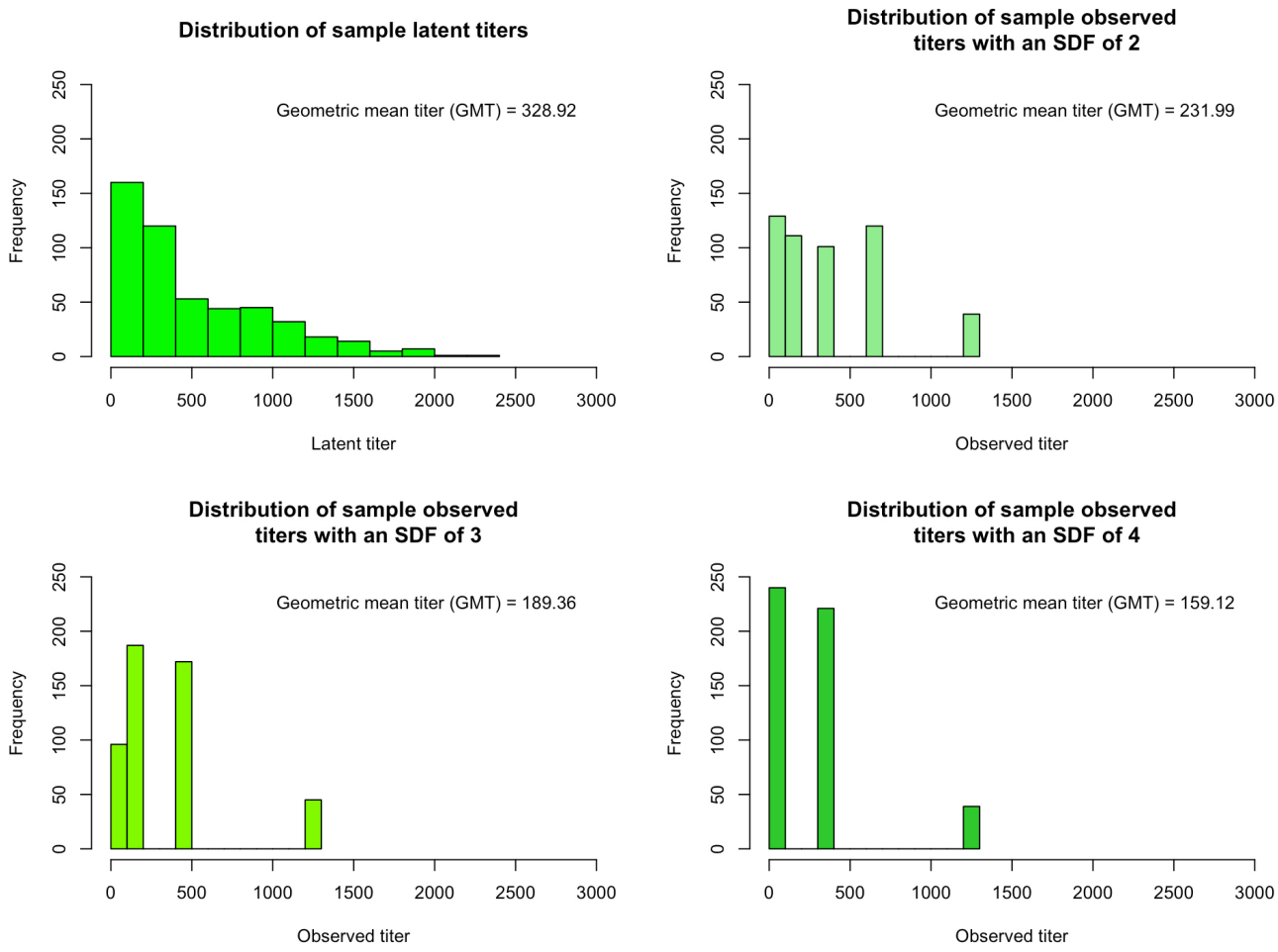

Figure 4: Simulated effect of discretizing titers, with increasing SDF

Another source of error that is a disadvantage to the HAI is the ULQ. As previously described, the HAI rounds down latent true titers to the closest observed dilution titer available. With a pre-selected ULQ set, any titers higher than the ULQ would thus be set to this observed upper limit no matter its magnitude. In the same way, the pre-selected starting dilution titer can also serve as a lower limit of quantification (LLQ), where titers lower than this starting titer would automatically be set to the next lowest observed titer regardless of magnitude. In both cases, they also lead to information loss. A higher ULQ and a lower starting titer would minimize this loss by allowing a greater range of titers to be observed in the HAI; however, this would require more dilution samples per patient, which in turn increases the cost, duration, processing time of the HAI. Resource availability or limitation during a clinical trial may compel setting a lower ULQ or higher starting titer, but the information loss in doing so must be considered before the analysis is conducted.

1.4 **Research Objective**

Hypothesis testing procedures suitable for two-group comparisons such as Pearson's χ^2 test or the Wilcoxon rank-sum test have been utilized to evaluate vaccine immunogenicity in prior studies involving the HAI [1, 15]. However, regression modeling has several added advantages compared to hypothesis testing, including being able to estimate the magnitude of vaccine immunogenicity [2]. Many of the hypothesis tests that are typically used are either exact or close approximations to corresponding saturated regression models; for example, the Wilcoxon rank-sum test is characteristic of a proportional odds model for comparing two groups [2], and a two-sample *t*-test of independent groups is equivalent to a simple linear regression with a binary exposure variable. Regression modeling is a more flexible approach to evaluating vaccine immunogenicity as they accommodate the inclusion of pre-treatment covariates. Despite regression modeling being an appealing alternative to hypothesis testing, the extent of the effect of inherent measurement error within the HAI due to SDF and the ULQ on the analysis of vaccine immunogenicity has not been adequately explored. In particular, the information loss and bias resulting from titer discretization have not been well characterized, even in simple two-group comparisons.

This research aims to study the effect of bias and efficiency loss associated with inherent measurement error due to titer discretization in an HAI analysis for evaluating vaccine immunogenicity. A set of simulations that seek to mimic a simple clinical trial in assessing the effect of a vaccine intervention were conducted under various data generating mechanisms. In particular, the effects of SDF and ULQ on regression parameters were emphasized and the results were compared between latent and observed titers to examine any errors or discrepancies that may arise between the models. Furthermore, an in-depth analysis on the relationship between the SDF – especially higher factors – and various regression parameters was explored. Various regression approaches were also applied to a real-world data set from a Phase I vaccine immunogenicity trial of solid organ transplant (SOT) patients.

CHAPTER 2

METHODS

2.1 Regression Modeling

As previously noted, regression modeling can be used to estimate the magnitude of the immunogenic effect while taking into account different sources of variation in the outcome. Two commonly implemented regression models will be specifically considered in this study: simple linear regression and proportional odds regression. A brief outline of each regression model is given below.

2.1.1 Simple Linear Regression

Consider a simple linear regression model in which the only predictor of interest is a patient's vaccination status. Let i index the independently sampled study subjects and $Y_i^{\bullet}$ be the latent true titer value of patient i. The observed titer of patient i, y_i, can be expressed as $y_i = \lfloor Y_i^{\bullet} \rfloor$, where $\lfloor \rfloor$ denotes the floor function. A simple linear regression model based on the observed titers can be written as [13]:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i,\ i = 1, 2, \ldots, n \tag{1}$$

where:

- y_i is the serum HAI titer of patient i
- β_0 is the mean serum HAI titer among patients who received the control
- β_1 is the difference in mean y_i between patients who received the experimental vaccine and patients who received the control
- x_i is the vaccination status of patient i
  - $x_i = 0$ indicates that patient i received the control
  - $x_i = 1$ indicates that patient i received the vaccine
- ϵ_i is the error term for patient i

Writing (1) for each of the n patients would yield the following n equations:

$$\begin{aligned} y_1 &= \beta_0 + \beta_1 x_1 + \epsilon_1 \\ y_2 &= \beta_0 + \beta_1 x_2 + \epsilon_2 \\ &\vdots \\ y_n &= \beta_0 + \beta_1 x_n + \epsilon_n \end{aligned}$$

The n equations above can also be written in matrix form as illustrated below:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}$$

which can be simplified to the following equation for simple linear regression [14]:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2}$$

Thus, $\mathbf{y}$ is denoted as the outcome vector of patient serum HAI titers, $\mathbf{X}$ is the design matrix of the patient vaccination status, $\boldsymbol{\beta}$ is the vector of regression coefficients, and $\boldsymbol{\epsilon}$ is the error vector. With this model, there are additional assumptions in order to make this model complete [13]:

- $\mathbb{E}(\boldsymbol{\epsilon}) = 0$ or $\mathbb{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$. This states that the regression model (2) is correct, in that $\mathbf{y}$ only depends on $\mathbf{X}$ and that any variation in $\mathbf{y}$ is random. When $\mathbf{X}$ is binary, the model is trivially satisfied as it is saturated. The model is often written in the expectation form:

$$\mathbb{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} \tag{3}$$

- $\mathrm{Cov}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{I}$ or $\mathrm{Cov}(\mathbf{y}) = \sigma^2\mathbf{I}$. This implies that the variance, σ^2, of $\boldsymbol{\epsilon}$ or of $\mathbf{y}$ does not depend on $\mathbf{X}$, and that each of the values in $\boldsymbol{\epsilon}$ or in $\mathbf{y}$ do not correlate with each other. This is known as homoscedasticity, or the assumption of equal variances. However, in many cases of real-life data, the assumption of homoscedasticity does not hold; heteroscedasticity is often implied instead, with the Huber-White estimator used to estimate the variance.
- Each of the patient observations are independent of each other.

With the observations in $\mathbf{y}$ and $\mathbf{X}$, the estimates for $\boldsymbol{\beta}$ and for the variance-covariance matrix of $\boldsymbol{\beta}$, $\mathrm{Var}(\boldsymbol{\beta})$, can be estimated. Let the estimate of $\boldsymbol{\beta}$ be represented as $\hat{\boldsymbol{\beta}}$, the estimates of β_0 and β_1 in $\boldsymbol{\beta}$ be denoted as $\hat{\beta}_0$ and $\hat{\beta}_1$ respectively, and the variance-covariance matrix for the estimate of $\boldsymbol{\beta}$ be $\mathrm{Var}\left(\hat{\boldsymbol{\beta}}\right)$.

The method of least squares is the best approach to obtain $\hat{\boldsymbol{\beta}}$ as it does not rely on any distributional assumptions, and the estimates of $\hat{\boldsymbol{\beta}}$ are unbiased [13]. This method evaluates the estimates by minimizing the sum of squares of the deviations of each y_i from their predicted values, denoted as $\hat{y}_i$, where $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$. The equation to estimate $\hat{\boldsymbol{\beta}}$ using the method of least squares for simple linear regression is:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}' \tag{4}$$

$\text{Var}\left(\hat{\boldsymbol{\beta}}\right)$ is evaluated from $\text{Cov}\left(\hat{\boldsymbol{\beta}}\right)$, or the estimated variance-covariance matrix of $\boldsymbol{\beta}$. Under the assumption that $\text{Cov}(\mathbf{y}) = \sigma^2\mathbf{I}$, $\text{Cov}\left(\hat{\boldsymbol{\beta}}\right) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. As mentioned previously, $\text{Cov}(\mathbf{y}) \neq \sigma^2\mathbf{I}$ in most cases and the actual variance-covariance matrix is not easily evaluated. Let $\text{Cov}(\mathbf{y}) = \mathbf{D}$, where $\mathbf{D}$ is the diagonal matrix of the squared residuals. $\text{Cov}\left(\hat{\boldsymbol{\beta}}\right)$ is estimated using the Huber-White estimator:

$$\text{Cov}\left(\hat{\boldsymbol{\beta}}\right) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \tag{5}$$

The variance and standard error of $\hat{\boldsymbol{\beta}}$, represented as $\text{Var}\left(\hat{\boldsymbol{\beta}}\right)$ and $\sigma\left(\hat{\beta}_1\right)$ respectively, can thus be obtained, where $\text{Var}\left(\hat{\boldsymbol{\beta}}\right)$ is the diagonal matrix of $\text{Cov}\left(\hat{\boldsymbol{\beta}}\right)$, and $\sigma\left(\hat{\beta}_1\right) = \sqrt{\text{Var}\left(\hat{\boldsymbol{\beta}}\right)}$. Using squared residuals instead of setting a particular covariance structure for $\mathbf{y}$ would ensure that the error structure is not grossly misspecified, as the covariance structure of the errors is the same as that of the outcome.

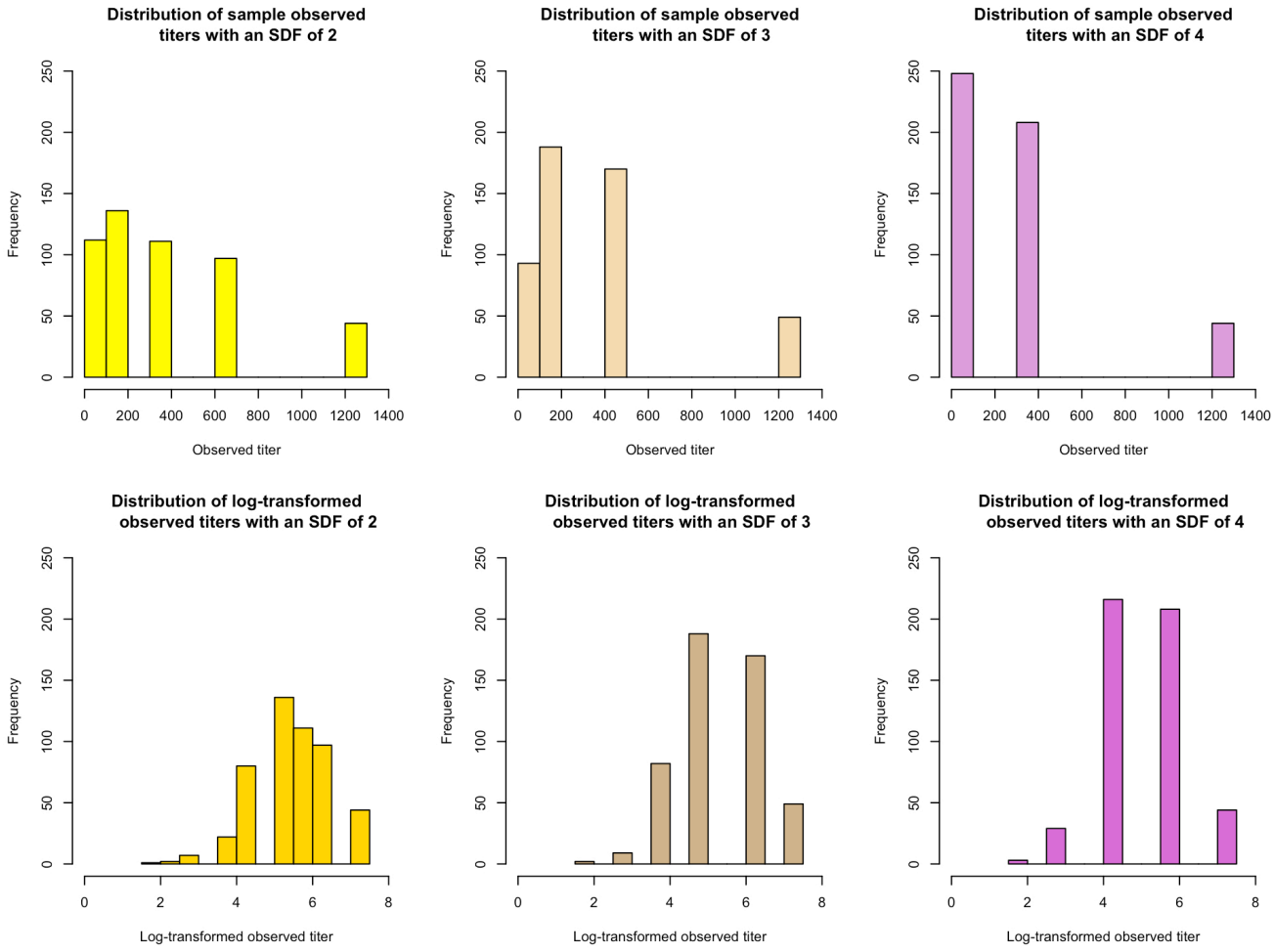

Figure 5: Visual distribution of untransformed and log-transformed titers under varying SDFs

It is often common practice for HAI titers to be log-transformed before the linear regression analysis is performed [10, 15]. Figure 5 in the previous page highlights the differences in the distribution of both untransformed and log-transformed titers. As subsequent HAI titers are multiplicative in powers of the SDF, log-transforming the titers would change the relationship between titer values to an additive association. The additive association induced by the log-transformation would thus be more applicable for use in simple linear regression modeling. Thus, from (3), the simple linear regression model under the log-transformation of the outcome is written in expectation form as:

$$\mathbb{E}(\log \mathbf{y}) = \mathbf{X}\boldsymbol{\beta} \tag{6}$$

or as follows:

$$\mathbb{E}\left(\log y_i\right) = \beta_0 + \beta_1 x_i,\ i = 1, 2, \ldots, n \tag{7}$$

From (7), exponentiating β_0 and β_1 to invert the scale change of the log-transformation leads to more applicable interpretations; e^{β_0} denotes the GMT of patients who received the control and e^{β_1} is the GMTR when comparing patients who received the vaccine to those who received the control. The practice of log-transformation in regression modeling also heavily corresponds to the frequent use of GMT and GMTR as the statistical measure in evaluating vaccine immunogenicity [1, 5, 6, 10, 11, 15].

2.1.2 Proportional Odds Regression

An alternative regression model that can also be considered is the proportional odds model, a type of cumulative probability model (CPM) that was first developed as an extension of logistic regression to ordered categorical data [7]. Due to the discretization of latent titers into observed titer levels, CPMs are a viable alternative to model the data; they can still be useful even if more refined SDFs are used [7]. In the case of HAI titers, higher titers are more indicative of immunogenic responses and are thus more desirable. In addition, titer categorization is already underscored in the HAI procedure and analysis from the discretization of latent titers.

To model a proportional odds regression for HAI titers, let Y_i be the outcome level and $\mathbf{Y}_i$ be the vector of outcome level indicators for patient i. Y_i can take K distinct numerical values that correspond to the outcome level, where K is the total number of distinct observed titer categories that is dependent on the SDF and ULQ; thus, $Y_i \in \{1, 2, \ldots, K\}$. $\mathbf{Y}_i$ is of length K, and $\mathbf{Y}_i = \left(Y_{i1}, Y_{i2}, \ldots, Y_{iK}\right)$, where $Y_{ik} = 1$ indicates that the observed titer for patient i is level k, and 0 otherwise. Thus, for any patient i,

all the elements in $\mathbf{Y}_i$ except for the observed level would be 0. Given n patients in a clinical trial, for the i^{th} patient,

$$\mathbf{Y}_i \sim \text{Multinomial}\left(1, \boldsymbol{\pi}_i\right)$$

where $\boldsymbol{\pi}_i = \left(\pi_{i1}, \pi_{i2}, \ldots, \pi_{iK}\right)$, $\pi_{ik} = P\left(Y_{ik} = 1\right)$, and $\sum_{k=1}^{K} \pi_{ik} = 1$.

For all patients, let $\pi_k(\mathbf{x}) = P\left(Y = k \mid \mathbf{X} = \mathbf{x}\right)$ and $\gamma_k(\mathbf{x}) = P\left(Y \leq k \mid \mathbf{X} = \mathbf{x}\right)$, where $\pi_k(\mathbf{x})$ is the probability that Y is equal to the outcome level k given the predictor vector $\mathbf{x}$, and $\gamma_k(\mathbf{x})$ is the cumulative probability that Y is at most at outcome level k given the predictor vector $\mathbf{x}$. From these definitions, $\gamma_k(\mathbf{x})$ can be evaluated for all values of k as follows:

$$\begin{aligned}
\gamma_1(\mathbf{x}) &= \pi_1(\mathbf{x}) \\
\gamma_2(\mathbf{x}) &= \pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) \\
&\vdots \\
\gamma_{K-1}(\mathbf{x}) &= \pi_{k_1}(\mathbf{x}) + \ldots + \pi_{K-1}(\mathbf{x}) \\
\gamma_K(\mathbf{x}) &= \pi_1(\mathbf{x}) + \ldots + \pi_K(\mathbf{x})
\end{aligned}$$

The cumulative logit is then defined as $\text{logit}\left[\gamma_k(\mathbf{x})\right] = \log\left[\dfrac{\gamma_k(\mathbf{x})}{1-\gamma_k(\mathbf{x})}\right] = \log\left[\dfrac{P\left(Y \leq k \mid \mathbf{X} = x\right)}{P\left(Y > k \mid \mathbf{X} = \mathbf{x}\right)}\right]$,

which is the log-odds of being at or below a response level of k. With K distinct values for Y, there would be K cumulative logit regression models, where for $k = 1, 2, \ldots, K$,

$$\log\left[\frac{\gamma_k(\mathbf{x})}{1-\gamma_k(\mathbf{x})}\right] = \boldsymbol{\beta}_{k0} + \boldsymbol{\beta}_{k1}\mathbf{x}$$

where $\boldsymbol{\beta}_{k0}$ is the log-cumulative odds of being at or below an outcome level k for the control (untreated) group, and $\boldsymbol{\beta}_{k1}$ is the difference in log-cumulative odds of being at or below an outcome level k when comparing the vaccine (treated) group to the control group. When exponentiated, $e^{\boldsymbol{\beta}_{k0}}$ becomes the cumulative odds of being at or below an outcome level k for the control group, and $e^{\boldsymbol{\beta}_{k1}}$ is the cumulative odds ratio of being at or below an outcome level k when comparing the vaccine group to the control group. With the proportional odds model, there is the added assumption that the effect of $\mathbf{x}$ is considered to be constant across all K models. The proportional odds model is thus represented by:

$$\log\left[\frac{\gamma_k(\mathbf{x})}{1-\gamma_k(\mathbf{x})}\right] = \boldsymbol{\beta}_{k0} + \boldsymbol{\beta}_1\mathbf{x} \tag{8}$$

where $\boldsymbol{\beta}_1$ is the fixed difference in log-cumulative odds of being at or below a response level k when comparing the vaccine group to the control group, and $e^{\boldsymbol{\beta}_1}$ is the fixed cumulative odds ratio of being at or below an outcome level k when comparing the vaccine group to the control group.

2.2 **Simulation**

All simulations were conducted using R version 4.0.5 and RStudio version 1.2.1335. A summary of the R program and the accompanying packages used for this research can be found in Appendix E. The following parameters for each simulation study were set for the data generating mechanism under the assumption of a simulation of a simple vaccine immunogenicity clinical trial for an experimental vaccine using two treatment groups.

2.2.1 Regression Modeling

Let every subject in this hypothetical clinical trial have a latent true serum HAI titer; the HAI titer is considered to be continuous along the positive real number line. The latent titers for both the control (untreated) and the vaccine (treated) groups were simulated under a Weibull distribution. The Weibull distribution is used as the data generating distribution as it is highly flexible in modeling symmetric or skewed data; titers are often positive- or right-skewed as the expected frequencies of people with higher titers decreases with increasing titer values. In addition, the support of the Weibull distribution is the positive real numbers, similar to the nature of HAI titers. The shape and scale of the Weibull distribution from which the data will be generated was determined by:

- Level of desired data skewness (low and high) for the control group
- Level of desired data skewness (low and high) and the treatment effect (none, trivial, moderate, and high) for the vaccine group

The latent titer distributions used to generate the data for each treatment group are based on a prior research by Nhat et al. that studied the general-population antibody titer distributions to the influenza A virus; the typical range of most log-transformed titers is from 1 to 7, translating to titer values between 2.7 and 1096.6 [10]. The treatment effect is varied using the Weibull scale parameter and the skewness is modified using both the Weibull shape and scale parameters; in this simulation, the shape parameter under high skewness is lower than that under low skewness. Figures 6 and 7 show the

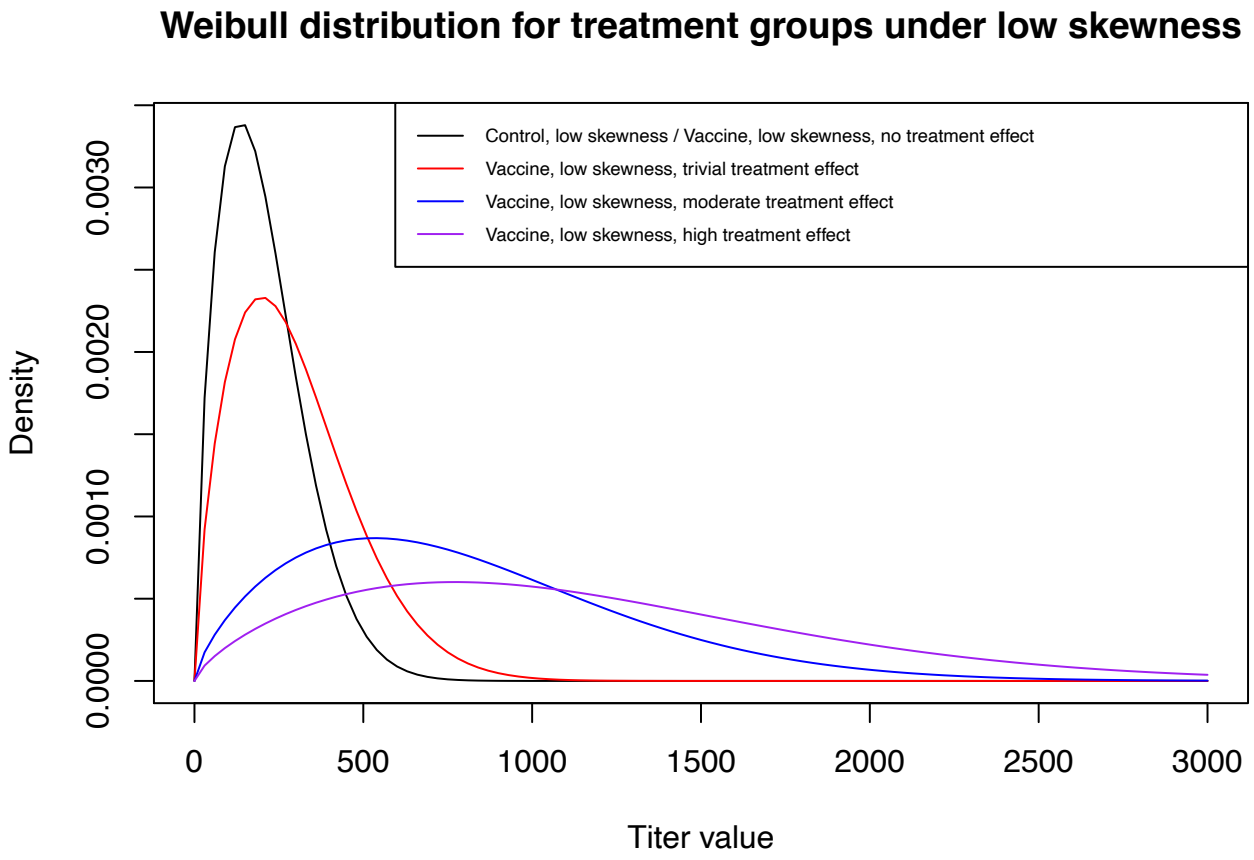

Figure 6: Distribution of simulated latent titers at low skewness via a Weibull data generating mechanism under different settings of treatment group and treatment effect

distribution of generated latent titers for several parameter combinations under low and high skewness respectively; high skewness in this study refer to greater right-skewness, with more titers concentrated towards lower values than that of low skewness. It is important to note that the Weibull shape and scale parameters used in this simulation are only one example of the many possibilities to simulate the titer data.

Examining the latent titer distributions under low skewness relative to the black line – either the control group or the vaccine group with the assumption of no treatment effect for the experimental vaccine – in Figure 6, with increasing treatment effect for the vaccine group, it can be observed that the proportions of subjects with higher latent titer values also increases. This is to conform to the expectation that a greater treatment effect from the vaccine being tested would lead to higher proportions of patients with larger observed titers, derived from the latent titers.

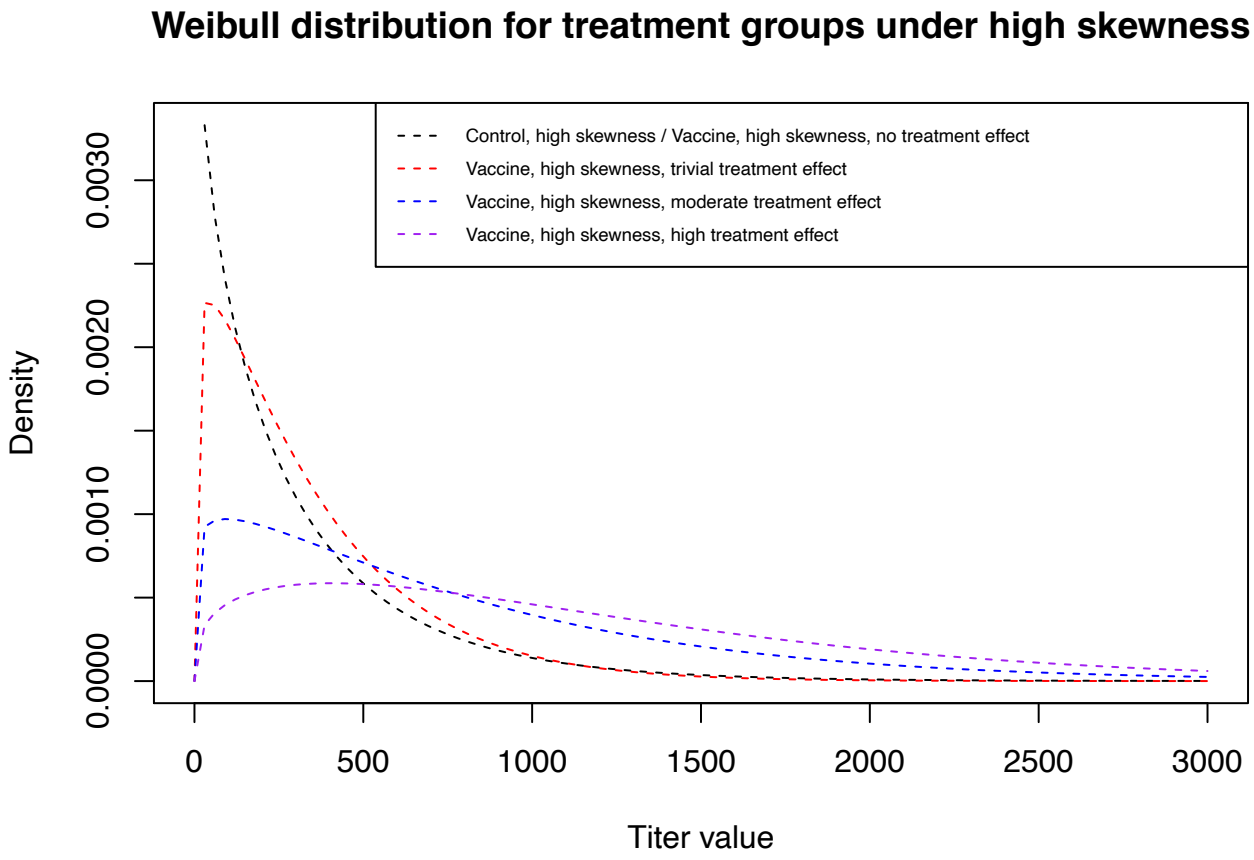

Figure 7: Distribution of simulated latent titers at high skewness via a Weibull data generating mechanism under different settings of treatment group and treatment effect

On the other hand, in observing the titer distributions under low skewness relative to the black line in Figure 7, the same trends can be concluded, but each of the plots is more heavily right-skewed compared to their corresponding counterparts of the same color as in Figure 6. This setting for high skewness is to reflect real-life circumstances that a higher proportion of people tend to have lower titers in general within a population due to the mechanisms of the innate immune system.

In addition to simulating latent true titers, a simulation of the "true population difference" in log-transformed titers was also included in order to set a hypothetical true treatment effect that will be used to compare the biases, if any, that are induced in the estimated treatment effect using both the latent and observed titers. The data for the treatment groups are generated and the observed titers are subsequently determined from the default SDF of 2 in keeping with common analytical practice, the starting dilution titer, and a pre-selected ULQ based on the simulation setting. In fitting a simple linear regression model to the simulated data to evaluate vaccine immunogenicity, $\hat{\beta}_1$ is evaluated from the

result derived from Equation (4) and $\sigma\left(\hat{\beta}_1\right)$ is obtained from the second entry of the result in Equation (9) that is derived from Equation (5):

$$\sigma\left(\hat{\boldsymbol{\beta}}\right) = \sqrt{\text{Var}\left(\hat{\boldsymbol{\beta}}\right)} = \sqrt{\text{diag}\left[\text{Cov}\left(\hat{\boldsymbol{\beta}}\right)\right]} \tag{9}$$

Each simulation was repeated 5000 times. From the results of all the trials, key statistics of interest were evaluated for both the latent and observed titer values. The simulations were then repeated with decreasing SDFs from 2 to a minimum of 1.1, and the inclusion of an increased ULQ that is greater than the highest possible latent titer that can be simulated. Based on the results obtained from prior simulations, the simulation parameters to be varied were selected in order to explore salient effects of the SDF as well as that of both a restricted and unrestricted ULQ on the regression results.

The data generation and model fitting were also repeated using the proportional odds regression, with 1000 repetitions for each simulation performed in this analysis to obtain the same statistics of interest above. Apart from the coverage probability of the treatment parameter and the power, the following table summarizes the additional key statistics evaluated for each simulation type where applicable, with different interpretations based on the type of regression modeled:

Statistic	Simple linear regression	Proportional odds regression
$\overline{\hat{\beta}}_1$	Mean log-transformed treatment estimate	Log-cumulative odds of the treatment parameter
$\sigma_e\left(\hat{\beta}_1\right)$	Estimated standard deviation of the log-transformed treatment estimate	Estimated standard deviation of the log-cumulative odds of the treatment parameter
$\overline{\sigma}\left(\hat{\beta}_1\right)$	Average standard error of the log-transformed treatment estimate	Average standard error of the log-cumulative odds of the treatment parameter

Table 2: Description of key statistics of interest based on regression type

2.2.2 SDF Association to Regression Statistics

The simulations for the regression analyses described in Section 2.2.1 only consider SDFs between 1.1 and 2, as higher SDFs are very rarely used in conventional analyses. While the impact of SDF on the HAI measurement error among common SDF values have not been explored, the overall association between the SDF – including values greater than 4 – and various regression parameter statistics is much less so. This presents an opportunity to conduct further simulation studies to

determine the extent of the effect of a greater range of SDFs on the parameter statistics obtained from both the simple linear regression and proportional odds regression analyses fitted to generated data described in the previous section. The effects of three main parameters will be studied:

- SDF, ranging from 1.1 to 10 in 0.1 increments
- True treatment effect of the vaccine, designated as β_1 in regression modeling. Values of β_1 to be considered are log(1), log(1.1), log(1.4), log(1.8), and log(3); these represent the log-transformation of typical GMTRs obtained from previous immunogenicity studies.
- Shape parameter of the Weibull data generating distribution, denoted as λ. As λ influences the slope of the Weibull distribution, it also indirectly affects the level of skewness of the distribution. Values of λ to be considered are 1.2, 1.5, 2.0, and 2.5.

This simulation, like that of regression modeling, mimics that of a vaccine immunogenicity clinical trial for an experimental vaccine using two treatment groups. With n = 500 patients enrolled into each treatment group, latent titer data for each patient were simulated from a Weibull data generating distribution and are then discretized to obtain the observed HAI titers. Each different simulation comprises a different combination of SDF, β_1, and λ which would affect the data generation and observed titer categories. Both a simple linear regression model and a proportional odds regression model were fitted to the observed titers, and the data generation and model fitting are repeated 500 times. In addition to the key statistics outlined in Table 2, another statistic of interest was evaluated from the repeated simulation: $e^{\bar{\hat{\beta}}_1}$, which represents the GMTR of the treatment estimate for simple linear regression, or the cumulative odds ratio for proportional odds regression. This procedure is repeated for each unique SDF, β_1, and λ combination. For each statistic of interest in both regression models, the data were plotted against the SDF, stratified by λ and β_1. The plots were then observed and analyzed to determine any patterns or interesting relationships.

CHAPTER 3

RESULTS

The salient results of all simulations for both regression model types are presented in this section. Additional results from each simulation conducted for each different setting can be found in the Appendices where indicated.

To calculate the precision level of the coverage probability (CP) of the treatment parameter, the standard error of the Bernoulli distribution will be used as the expectation of the Bernoulli distribution is itself a proportion. The precision level of the CP would thus be: $\sigma_{CP} = \sqrt{\frac{(1-\alpha)\alpha}{n_S}}$, where:

- σ_{CP} is the precision, i.e., standard error of the CP
- α is the significance level, i.e., Type I error rate
- n_S is the number of simulation repetitions.

In this study, $\alpha = 0.05$.

3.1 Simple Linear Regression

Each of the following tables displays the summarized results of the 5000 repeated trials for each simple linear regression simulation performed in the following order of parameter settings, with a stratification either by treatment effect, sample sizes for each treatment group, or both:

- Default SDF of 2 and restricted ULQ of 1280; stratified by treatment effect
- Decreasing SDFs up to 1.1 and restricted ULQ based on SDF; stratified by sample size
- Decreasing SDFs up to 1.1 and unrestricted ULQ; stratified by treatment effect and sample size

Each subsequent simulation as above was conducted with the parameter settings selected based on the results obtained from the previous simulation study. The standard error of the CP obtained for all simple linear regression models fitted to the data is $\sqrt{\frac{0.95 \cdot 0.05}{5000}} \approx 0.003$. This translates to an overall precision of the CP of ±0.01.

3.1.1 Restricted ULQ and Default SDF

[1]*Large treatment effect*

[1]CS	TS	SS	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**50**	1.719	1.719	0.148	0.146	0.943	1.000	1.689	0.150	0.147	0.941	1.000	0.976
Low	**Low**	**250**	1.719	1.719	0.067	0.066	0.949	1.000	1.688	0.067	0.067	0.925	1.000	0.981
Low	**Low**	**500**	1.717	1.716	0.047	0.047	0.949	1.000	1.686	0.047	0.047	0.898	1.000	0.994
Low	**High**	**50**	1.579	1.577	0.175	0.170	0.942	1.000	1.520	0.171	0.167	0.934	1.000	1.049
Low	**High**	**250**	1.581	1.581	0.076	0.077	0.949	1.000	1.523	0.075	0.075	0.889	1.000	1.043
Low	**High**	**500**	1.581	1.580	0.055	0.054	0.948	1.000	1.522	0.054	0.053	0.805	1.000	1.037
High	**Low**	**50**	1.723	1.720	0.217	0.212	0.939	1.000	1.678	0.213	0.208	0.931	1.000	1.036
High	**Low**	**250**	1.725	1.724	0.096	0.096	0.948	1.000	1.682	0.093	0.094	0.925	1.000	1.047
High	**Low**	**500**	1.725	1.725	0.069	0.068	0.945	1.000	1.683	0.068	0.067	0.900	1.000	1.028
High	**High**	**50**	1.587	1.583	0.231	0.229	0.945	1.000	1.515	0.225	0.222	0.931	1.000	1.060
High	**High**	**250**	1.588	1.585	0.104	0.104	0.949	1.000	1.516	0.102	0.100	0.884	1.000	1.057
High	**High**	**500**	1.588	1.588	0.074	0.073	0.948	1.000	1.519	0.072	0.072	0.835	1.000	1.064

Table 3: Simulation results under a large treatment effect

Under a large treatment effect, there is a small but noticeable bias in MTE for the model using observed titers compared to that with latent titers. In addition, the model with latent titers seems to be more efficient when both treatment groups have low skewness, whereas the efficiency shifts slightly in favor of the model with observed titers when either group has a highly skewed distribution. The ESD and ASE are also almost equal to each other, and both become smaller when the group sample sizes

[1] The full term for each statistical abbreviation in the headings of Table 3 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SS – Sample size per treatment group
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

increase. Both the ESD and ASE magnitudes are influenced by the group skewness combination. P obtained for both models equals 1 due to the high estimated GMTR between $e^{1.094} \approx 4.9$ and $e^{1.725} \approx 5.6$, but the CP shows a decreasing trend for the model with observed titers when the sample sizes increase. The simulation results under no and moderate treatment effects for this setting can be found in Appendix A for comparison with the results in Table 3.

From this first simulation setting, when comparing the effect of different treatment levels on the regression parameters, there seems to be minor to no observable differences between the ESD and ASE across all treatment effect levels as they are almost equal not only to each other, but also across titer types. This is a good indication that observed titers do not inflate the variance of the treatment parameter. However, as the treatment effect increases, there is a slight but noticeable increase in the bias of MTE obtained from the model with observed titers. Neither the MTE bias nor the RE seem to change for each group skewness combination when the group sample sizes increase; this indicates that the sample sizes do not influence the reliability of the model. Furthermore, with the given α, the CP starts to deviate away from the expected proportion of 0.95 as the group sample sizes increase. On the other hand, P immediately trends towards 1 even under a moderate treatment effect with obtained estimated GMTRs of at least 3.

The next simulation setting that includes decreasing SDFs focused exclusively on the large treatment effect due to the increased MTE bias and the decreasing CP found for this treatment effect level and not for those under no and moderate treatment effects. It is of interest to examine whether the MTE and CP estimates improve as the SDF becomes more granular, in addition to evaluating other significant changes to the simple linear regression results. The following table summarizes the total distinct observed titer levels as well as the ULQ to each SDF for the next simulation setting. ULQ for SDFs that are less than 2 were set to the titer with the smallest difference to the ULQ for the default SDF.

SDF	2.0	1.9	1.8	1.7	1.6	1.5	1.4	1.3	1.2	1.1
Total distinct titer levels	8	9	10	11	12	14	17	22	31	59
Maximum dilution titer	1280	1613	1785	1714	1407	1460	1525	1606	1424	1384

Table 4: Total distinct titer levels and modified restricted ULQ to each decreasing SDF

3.1.2 Restricted ULQ, Large Treatment Effect, and Decreasing SDF

[2]*Large sample sizes* (*500 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	1.721	1.717	0.046	0.047	0.950	1.000	1.687	0.047	0.047	0.890	1.000	0.985
Low	**Low**	**1.8**	1.719	1.717	0.047	0.047	0.948	1.000	1.711	0.048	0.048	0.945	1.000	0.965
Low	**Low**	**1.6**	1.716	1.718	0.047	0.047	0.944	1.000	1.679	0.047	0.046	0.872	1.000	1.026
Low	**Low**	**1.4**	1.719	1.718	0.047	0.047	0.950	1.000	1.679	0.045	0.046	0.857	1.000	1.048
Low	**Low**	**1.2**	1.718	1.718	0.047	0.047	0.950	1.000	1.646	0.045	0.044	0.642	1.000	1.112
Low	**High**	**2.0**	1.580	1.581	0.056	0.054	0.944	1.000	1.523	0.055	0.053	0.812	1.000	1.028
Low	**High**	**1.8**	1.581	1.581	0.054	0.054	0.956	1.000	1.561	0.055	0.054	0.935	1.000	0.992
Low	**High**	**1.6**	1.581	1.580	0.054	0.055	0.953	1.000	1.514	0.052	0.052	0.750	1.000	1.079
Low	**High**	**1.4**	1.579	1.580	0.055	0.054	0.945	1.000	1.516	0.052	0.052	0.771	1.000	1.097
Low	**High**	**1.2**	1.581	1.580	0.055	0.054	0.952	1.000	1.480	0.051	0.051	0.490	1.000	1.150
High	**Low**	**2.0**	1.725	1.723	0.068	0.068	0.944	1.000	1.681	0.067	0.067	0.895	1.000	1.031
High	**Low**	**1.8**	1.724	1.724	0.068	0.068	0.953	1.000	1.707	0.067	0.067	0.942	1.000	1.013
High	**Low**	**1.6**	1.725	1.725	0.069	0.068	0.947	1.000	1.677	0.067	0.066	0.881	1.000	1.054
High	**Low**	**1.4**	1.725	1.724	0.067	0.068	0.951	1.000	1.677	0.065	0.066	0.893	1.000	1.061
High	**Low**	**1.2**	1.723	1.724	0.068	0.068	0.949	1.000	1.647	0.066	0.065	0.778	1.000	1.082
High	**High**	**2.0**	1.588	1.587	0.075	0.073	0.945	1.000	1.518	0.073	0.071	0.829	1.000	1.054
High	**High**	**1.8**	1.585	1.585	0.074	0.073	0.952	1.000	1.554	0.072	0.072	0.927	1.000	1.039
High	**High**	**1.6**	1.586	1.589	0.073	0.073	0.947	1.000	1.514	0.070	0.071	0.825	1.000	1.079
High	**High**	**1.4**	1.583	1.586	0.072	0.073	0.952	1.000	1.513	0.069	0.070	0.836	1.000	1.084
High	**High**	**1.2**	1.588	1.585	0.074	0.073	0.947	1.000	1.480	0.070	0.070	0.648	1.000	1.113

Table 5: Simulation results under large sample sizes

[2] The full term for each statistical abbreviation in the headings of Table 5 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

The results in Table 5 show similar trends for the MTE bias, ESD, and ASE to that of Table 3, except that the ESD and ASE do not seem to be influenced by changes to the SDF. The RE shows a slow increasing trend as the SDF decreases, indicating smaller variances for the model with observed titers than that with latent titers. The CP shows a non-monotonic decreasing trend, with much faster declining rates at smaller SDFs. The simulation results under small and medium sample sizes for this setting can be found in Appendix B for comparison with the results in Table 5.

From this second simulation setting that includes decreasing SDFs, there seems to be trivial to no changes to the bias of the MTE when using observed titers as the group sample sizes increase. Most of the bias may be attributed to the differences in group skewness rather than the SDF. Nevertheless, the presence of a bias in the MTE highlights that observed titers do result in some loss of information under this setting.

At higher sample sizes, the CP suffers to a greater degree as the SDF decreases to much smaller factors; in some cases, it can even reach probability levels below 0.5. With such results that highlights the severity and the disadvantage of imposing a low ULQ (due to possible constraints on cost, time, and resources that would be required to conduct a HAI), it is of interest in examining whether increasing the ULQ beyond the highest possible titer generated by the simulation improves the CP as the SDF becomes more granular. In addition, the simulation was also used to examine the impact of trivial treatment effects in addition to large treatment effects on the regression results. Due to the massive increase in P when comparing a moderate treatment effect to no treatment effect, an additional treatment level was introduced to study this intermediate setting between the former and the latter.

The following table summarizes the total distinct observed titer levels as well as the unrestricted ULQ for each SDF, under the assumption that the highest possible titer generated by the data generating mechanism does not exceed this increased ULQ by at least 1 titer level. ULQ for SDFs that are less than 2 were set to the titer with the smallest difference to the ULQ for the default SDF.

SDF	2.0	1.9	1.8	1.7	1.6	1.5	1.4	1.3	1.2	1.1
Total distinct titer levels	10	11	12	13	15	17	21	27	38	73
Maximum dilution titer	5120	5825	5784	4952	5765	4926	5857	5963	5103	5256

Table 6: Total distinct titer levels and unrestricted ULQ to each decreasing SDF

3.1.3 Unrestricted ULQ and Decreasing SDF

i) [3]*Trivial treatment effect and small sample sizes* (*50 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	0.372	0.370	0.149	0.145	0.940	0.722	0.369	0.153	0.150	0.941	0.692	0.946
Low	**Low**	**1.8**	0.370	0.375	0.146	0.145	0.947	0.728	0.375	0.151	0.149	0.950	0.706	0.945
Low	**Low**	**1.6**	0.370	0.374	0.147	0.145	0.945	0.730	0.373	0.149	0.147	0.944	0.713	0.969
Low	**Low**	**1.4**	0.373	0.370	0.147	0.146	0.948	0.706	0.370	0.147	0.147	0.948	0.704	0.990
Low	**Low**	**1.2**	0.370	0.369	0.146	0.145	0.947	0.718	0.369	0.146	0.145	0.946	0.718	0.996
Low	**High**	**2.0**	0.256	0.256	0.188	0.185	0.946	0.302	0.259	0.190	0.188	0.947	0.296	0.985
Low	**High**	**1.8**	0.256	0.253	0.189	0.186	0.943	0.298	0.256	0.191	0.187	0.943	0.297	0.984
Low	**High**	**1.6**	0.255	0.255	0.189	0.186	0.942	0.297	0.258	0.189	0.187	0.943	0.293	1.001
Low	**High**	**1.4**	0.256	0.256	0.185	0.186	0.951	0.296	0.258	0.185	0.186	0.949	0.302	1.002
Low	**High**	**1.2**	0.255	0.253	0.187	0.186	0.948	0.291	0.255	0.186	0.185	0.948	0.297	1.010
High	**Low**	**2.0**	0.377	0.377	0.212	0.211	0.943	0.421	0.364	0.212	0.210	0.943	0.406	0.999
High	**Low**	**1.8**	0.376	0.374	0.212	0.211	0.947	0.426	0.363	0.211	0.210	0.944	0.409	1.010
High	**Low**	**1.6**	0.378	0.376	0.217	0.211	0.942	0.432	0.366	0.215	0.209	0.939	0.415	1.017
High	**Low**	**1.4**	0.376	0.375	0.216	0.211	0.946	0.418	0.366	0.213	0.208	0.947	0.410	1.024
High	**Low**	**1.2**	0.375	0.377	0.214	0.211	0.944	0.426	0.369	0.211	0.209	0.943	0.420	1.027
High	**High**	**2.0**	0.259	0.265	0.242	0.240	0.950	0.193	0.257	0.241	0.239	0.946	0.189	1.011
High	**High**	**1.8**	0.263	0.266	0.246	0.240	0.940	0.204	0.258	0.244	0.238	0.941	0.203	1.021
High	**High**	**1.6**	0.263	0.253	0.242	0.241	0.949	0.180	0.246	0.239	0.238	0.946	0.178	1.022
High	**High**	**1.4**	0.262	0.263	0.251	0.241	0.935	0.197	0.256	0.248	0.238	0.935	0.195	1.023
High	**High**	**1.2**	0.266	0.262	0.247	0.241	0.942	0.205	0.256	0.244	0.238	0.943	0.200	1.029

Table 7: Simulation results under small sample sizes

[3] The full term for each statistical abbreviation in the headings of Table 7 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

ii) [4]*Trivial treatment effect and medium sample sizes* (*250 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	0.370	0.372	0.066	0.066	0.948	1.000	0.371	0.068	0.068	0.948	0.999	0.944
Low	**Low**	**1.8**	0.369	0.371	0.066	0.066	0.944	1.000	0.371	0.068	0.067	0.944	1.000	0.954
Low	**Low**	**1.6**	0.370	0.369	0.066	0.066	0.954	1.000	0.369	0.067	0.067	0.953	1.000	0.967
Low	**Low**	**1.4**	0.371	0.372	0.065	0.066	0.948	1.000	0.372	0.066	0.066	0.951	1.000	0.981
Low	**Low**	**1.2**	0.368	0.371	0.065	0.066	0.952	1.000	0.371	0.065	0.066	0.952	1.000	0.996
Low	**High**	**2.0**	0.253	0.256	0.083	0.084	0.953	0.856	0.259	0.085	0.085	0.953	0.856	0.965
Low	**High**	**1.8**	0.257	0.257	0.085	0.084	0.946	0.850	0.260	0.086	0.085	0.946	0.854	0.986
Low	**High**	**1.6**	0.256	0.256	0.083	0.084	0.951	0.855	0.259	0.084	0.084	0.951	0.861	0.987
Low	**High**	**1.4**	0.255	0.257	0.084	0.084	0.951	0.863	0.259	0.084	0.084	0.951	0.869	1.003
Low	**High**	**1.2**	0.257	0.256	0.084	0.084	0.951	0.853	0.258	0.083	0.084	0.951	0.864	1.009
High	**Low**	**2.0**	0.377	0.377	0.096	0.095	0.949	0.978	0.363	0.096	0.095	0.945	0.971	1.013
High	**Low**	**1.8**	0.376	0.378	0.096	0.095	0.946	0.978	0.366	0.095	0.095	0.947	0.971	1.017
High	**Low**	**1.6**	0.376	0.376	0.095	0.095	0.953	0.982	0.366	0.095	0.094	0.948	0.977	1.013
High	**Low**	**1.4**	0.376	0.378	0.095	0.095	0.953	0.981	0.369	0.094	0.094	0.952	0.978	1.027
High	**Low**	**1.2**	0.379	0.376	0.095	0.095	0.947	0.979	0.368	0.094	0.094	0.947	0.977	1.022
High	**High**	**2.0**	0.262	0.261	0.111	0.109	0.944	0.664	0.253	0.110	0.108	0.941	0.638	1.011
High	**High**	**1.8**	0.264	0.261	0.109	0.109	0.950	0.666	0.253	0.108	0.108	0.949	0.650	1.021
High	**High**	**1.6**	0.261	0.263	0.109	0.109	0.951	0.678	0.256	0.108	0.108	0.949	0.658	1.018
High	**High**	**1.4**	0.261	0.266	0.109	0.109	0.950	0.686	0.260	0.108	0.108	0.951	0.674	1.024
High	**High**	**1.2**	0.264	0.260	0.109	0.109	0.947	0.663	0.255	0.108	0.108	0.949	0.654	1.026

Table 8: Simulation results under a trivial treatment effect and medium sample sizes

[4] The full term for each statistical abbreviation in the headings of Table 8 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

iii) [5]*Trivial treatment effect and large sample sizes* (*500 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	0.371	0.371	0.047	0.046	0.951	1.000	0.370	0.048	0.048	0.947	1.000	0.940
Low	**Low**	**1.8**	0.372	0.370	0.046	0.046	0.956	1.000	0.370	0.047	0.048	0.954	1.000	0.950
Low	**Low**	**1.6**	0.371	0.370	0.047	0.046	0.948	1.000	0.370	0.047	0.047	0.952	1.000	0.983
Low	**Low**	**1.4**	0.370	0.371	0.046	0.046	0.949	1.000	0.371	0.047	0.047	0.948	1.000	0.981
Low	**Low**	**1.2**	0.371	0.370	0.047	0.046	0.948	1.000	0.370	0.047	0.047	0.947	1.000	0.998
Low	**High**	**2.0**	0.256	0.258	0.059	0.059	0.955	0.992	0.261	0.059	0.060	0.953	0.992	0.980
Low	**High**	**1.8**	0.256	0.256	0.059	0.060	0.952	0.988	0.260	0.060	0.060	0.951	0.989	0.981
Low	**High**	**1.6**	0.256	0.256	0.060	0.059	0.945	0.987	0.259	0.061	0.060	0.944	0.989	0.996
Low	**High**	**1.4**	0.256	0.256	0.061	0.060	0.944	0.986	0.259	0.061	0.059	0.944	0.988	1.003
Low	**High**	**1.2**	0.257	0.257	0.059	0.060	0.953	0.991	0.259	0.059	0.059	0.953	0.993	1.008
High	**Low**	**2.0**	0.378	0.377	0.067	0.068	0.950	1.000	0.365	0.067	0.067	0.950	1.000	1.007
High	**Low**	**1.8**	0.378	0.377	0.068	0.068	0.948	1.000	0.366	0.068	0.067	0.948	1.000	1.017
High	**Low**	**1.6**	0.375	0.378	0.068	0.067	0.950	1.000	0.367	0.068	0.067	0.950	1.000	1.016
High	**Low**	**1.4**	0.374	0.379	0.067	0.068	0.951	1.000	0.370	0.066	0.067	0.951	1.000	1.024
High	**Low**	**1.2**	0.378	0.378	0.068	0.067	0.945	1.000	0.370	0.067	0.067	0.945	1.000	1.028
High	**High**	**2.0**	0.263	0.263	0.077	0.077	0.947	0.923	0.254	0.077	0.076	0.947	0.910	1.008
High	**High**	**1.8**	0.263	0.262	0.078	0.077	0.949	0.922	0.254	0.078	0.076	0.949	0.911	1.020
High	**High**	**1.6**	0.263	0.263	0.077	0.077	0.947	0.926	0.255	0.077	0.076	0.945	0.917	1.019
High	**High**	**1.4**	0.264	0.262	0.077	0.077	0.948	0.922	0.255	0.076	0.076	0.949	0.916	1.028
High	**High**	**1.2**	0.263	0.263	0.078	0.077	0.945	0.928	0.258	0.077	0.076	0.944	0.921	1.023

Table 9: Simulation results under a trivial treatment effect and large sample sizes

[5] The full term for each statistical abbreviation in the headings of Table 9 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

At small sample sizes, the RE increases as the SDF decreases, albeit at a slower rate compared to the setting with restricted ULQ. The MTE bias is heavily diminished with an unrestricted ULQ. The ESD and the ASE also show similar trends of being almost equal to each other. The CP is much closer to 0.95 across all parameters. P seems to vary depending on the group skewness combination, with a lower P obtained when the vaccine group is highly skewed. Similar trends can be observed for the RE, MTE, ESD, ASE, and CP as the group sample sizes increase. The ESD and the ASE show decreasing trends with increasing group sample sizes. P also increases substantially towards 1, while maintaining the same trend in differences between group skewness. The simulation results under a large treatment effect for this setting can be found in Appendix C for comparison with the results under corresponding sample sizes in Tables 7, 8, and 9.

Increasing the ULQ to be beyond that of the highest possible generated titer corrects the CP back to the expected probability of 0.95. There seems to be no observable or significant benefit to using smaller SDFs not only in this simulation setting but also in previous simulations that utilized them beyond a heavily improved CP; the MTE bias from the simulated TTP in the model using observed titers and that of the model with latent titers show barely any meaningful change as the SDF decreases, and the ESD and ASE parameters remained almost constant between titer types and across all combinations of group skewness. Without a loss of generality, it would be safe to assume that the same trends would be present if the simulation is conducted under both no and moderate treatment effects.

3.1.4 Extended SDF Impact on Regression Statistics

Though there is no discernible advantage in improving MTE, ESD or ASE estimates when applying SDFs smaller than 2 to a linear regression analysis of an HAI, there is still the possibility that higher SDFs can meaningfully impact the regression results and statistics. As stated in Section 1.3, SDFs higher than 2 have been used in previous research, and it cannot be ascertained that this uncommon practice would yield results with similar trends and estimated values as described in all the linear regression simulations.

The following figures in this section display the summarized simple linear regression results over 500 repeated trials for each simulation, of the four key statistics outlined in Sections 2.2.1 and 2.2.2 in evaluating the overall impact of a larger range of SDFs on these statistics of interest, stratified by both λ and β_1 that is defined under simple linear regression in Table 2.

i) *Mean log-transformed treatment estimate – MTE,* $\overline{\hat{\beta}}_1$

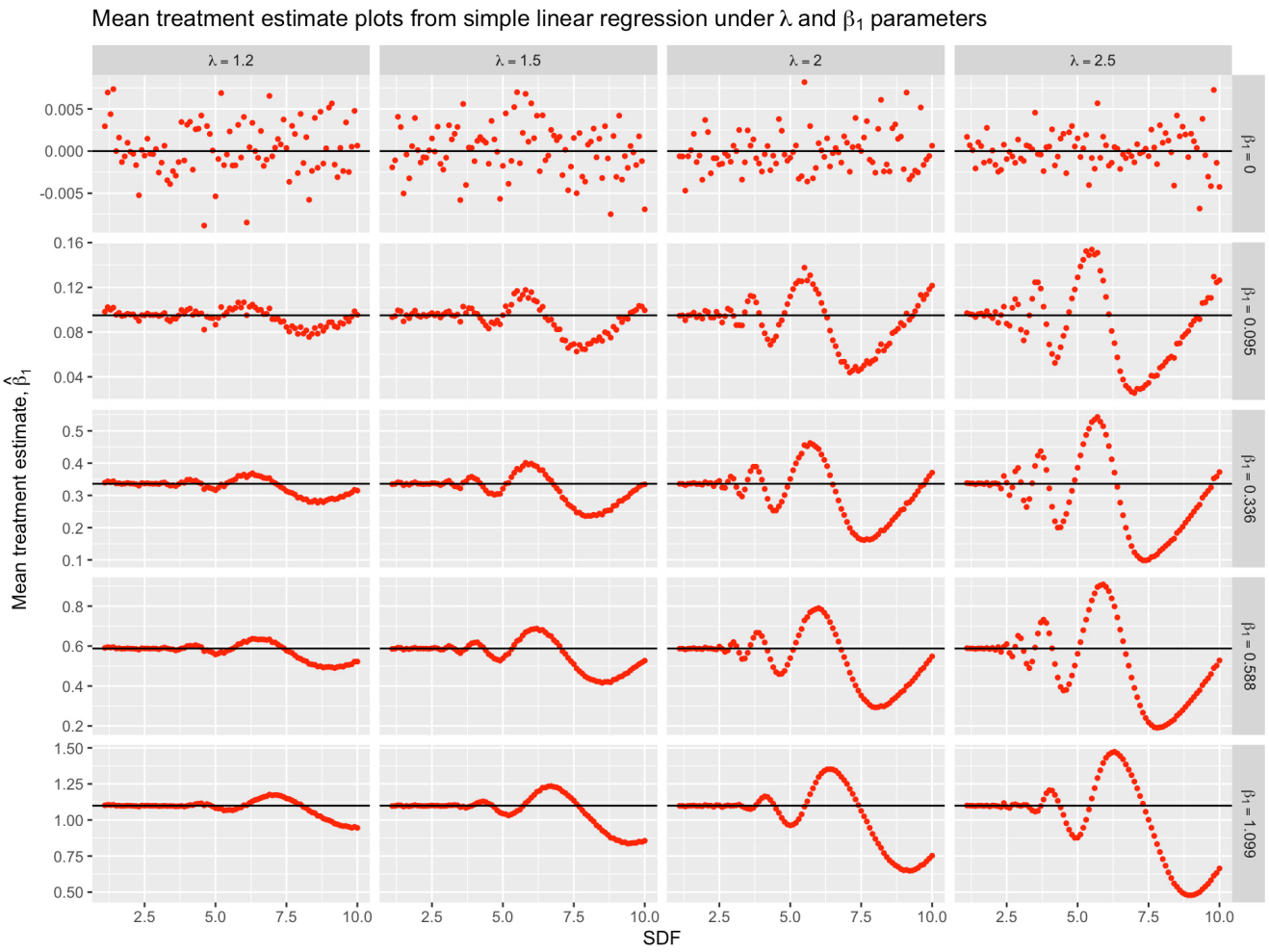

Figure 8: Scatter plots of MTEs from simple linear regression to SDFs under different λ and β_1

Figure 8 highlights an interesting relationship between the SDF and the MTE, $\overline{\hat{\beta}}_1$, when the range of the SDF is extended beyond 2. When $\beta_1 = 0$, there is little to no deviation of $\overline{\hat{\beta}}_1$ from the true effect of 0 regardless of the SDF utilized in the simulation. However, increasing λ in the Weibull distribution for the data generating mechanism seems to reduce the already small estimate biases closer to 0, thereby increasing the consistency and accuracy of the simulation. When $\beta_1 > 0$, the distribution of $\overline{\hat{\beta}}_1$ can be more clearly observed; there is an amplifying periodic relationship between $\overline{\hat{\beta}}_1$ and the SDF, the periodicity being only noticeable within a particular range of SDFs depending on both λ and β_1.

When comparing simulations with the same β_1 for $\beta_1 > 0$, increasing λ increases the periodic amplitude and decreases the periodic length. In addition, the minimum SDF in which the periodicity can begin to be observed decreases with the increase in λ. This results in an overall increased rate of the

bias of $\overline{\hat{\beta}}_1$ as the SDF increases. By comparing simulations with the same λ for $\beta_1 > 0$, it can be observed that increasing β_1 increases the precision of the periodic relationship; this can be observed as a gradual smoothing of the points in the periodic pattern as β_1 increases. In addition, the periodic length increases slightly, resulting in an increase to the minimum SDF in which the periodic pattern can begin to be observed. The bias in $\overline{\hat{\beta}}_1$ would thus not be as readily apparent in the statistical analysis within a greater range of lower SDFs. However, there are a few cases in which the bias of $\overline{\hat{\beta}}_1$ is either close to or exactly 0 for much higher SDFs due to this periodic nature; it would be difficult to identify these particular SDFs in practice without a prior simulation to test them. The increasing bias in MTE with larger SDFs are prominent through this simulation.

ii) *Geometric mean titer ratio of treatment estimate – GMTR,* $e^{\overline{\hat{\beta}}_1}$

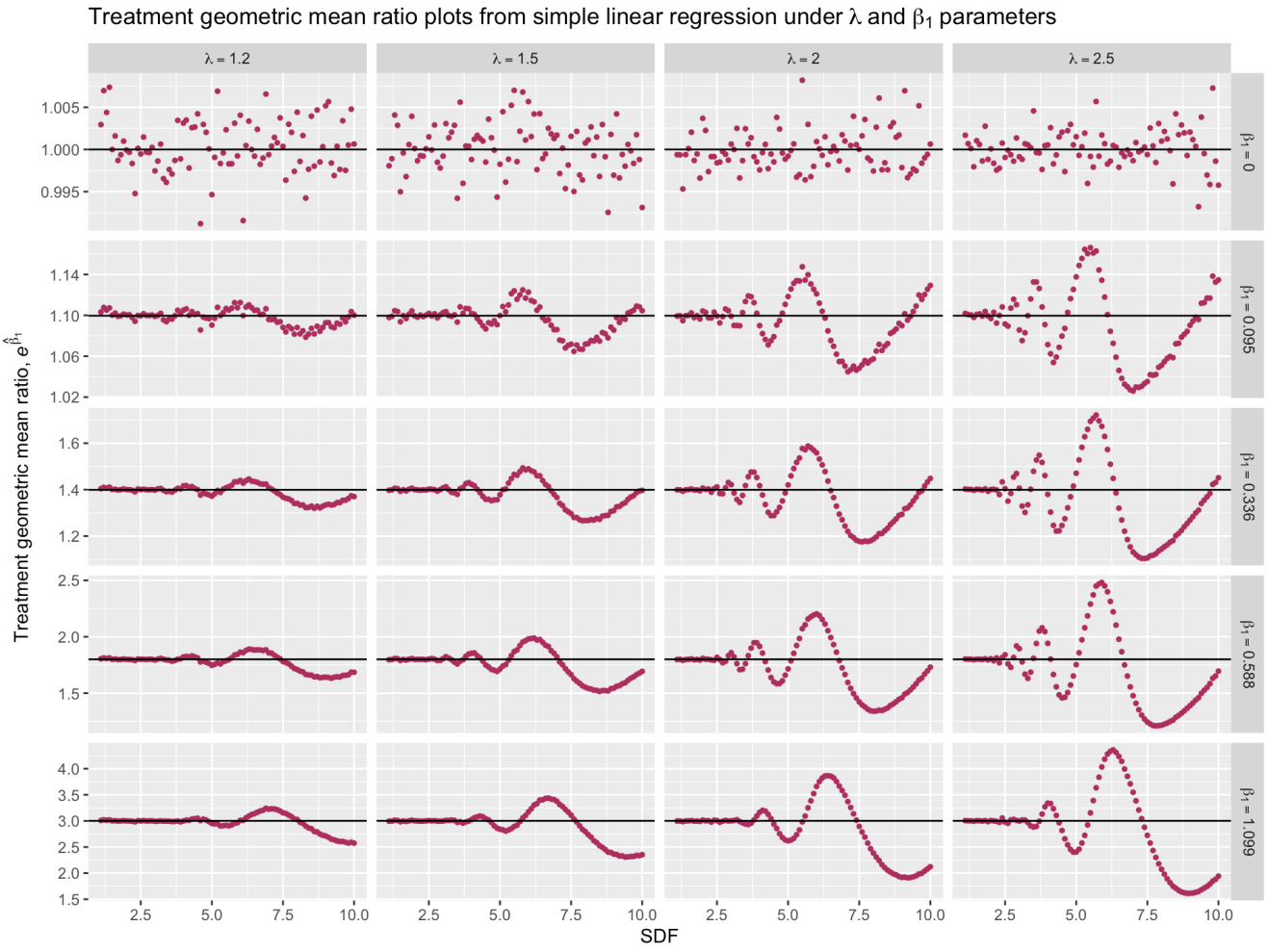

Figure 9: Scatter plots of GMTRs from simple linear regression to SDFs under different λ and β_1

The shapes and trajectories of the GMTR plots as shown in Figure 9 are exactly the same as that of the MTE as the GMTR is the exponentiation of $\overline{\hat{\beta}}_1$, a monotonic function. However, due to the exponentiation factor, a small change in the MTE would result in a much larger change in the GMTR as indicated by the scale of the *y*-axes of the plots.

iii) *Estimated standard deviation of the log-transformed treatment estimate – ESD,* $\sigma_e\left(\hat{\beta}_1\right)$

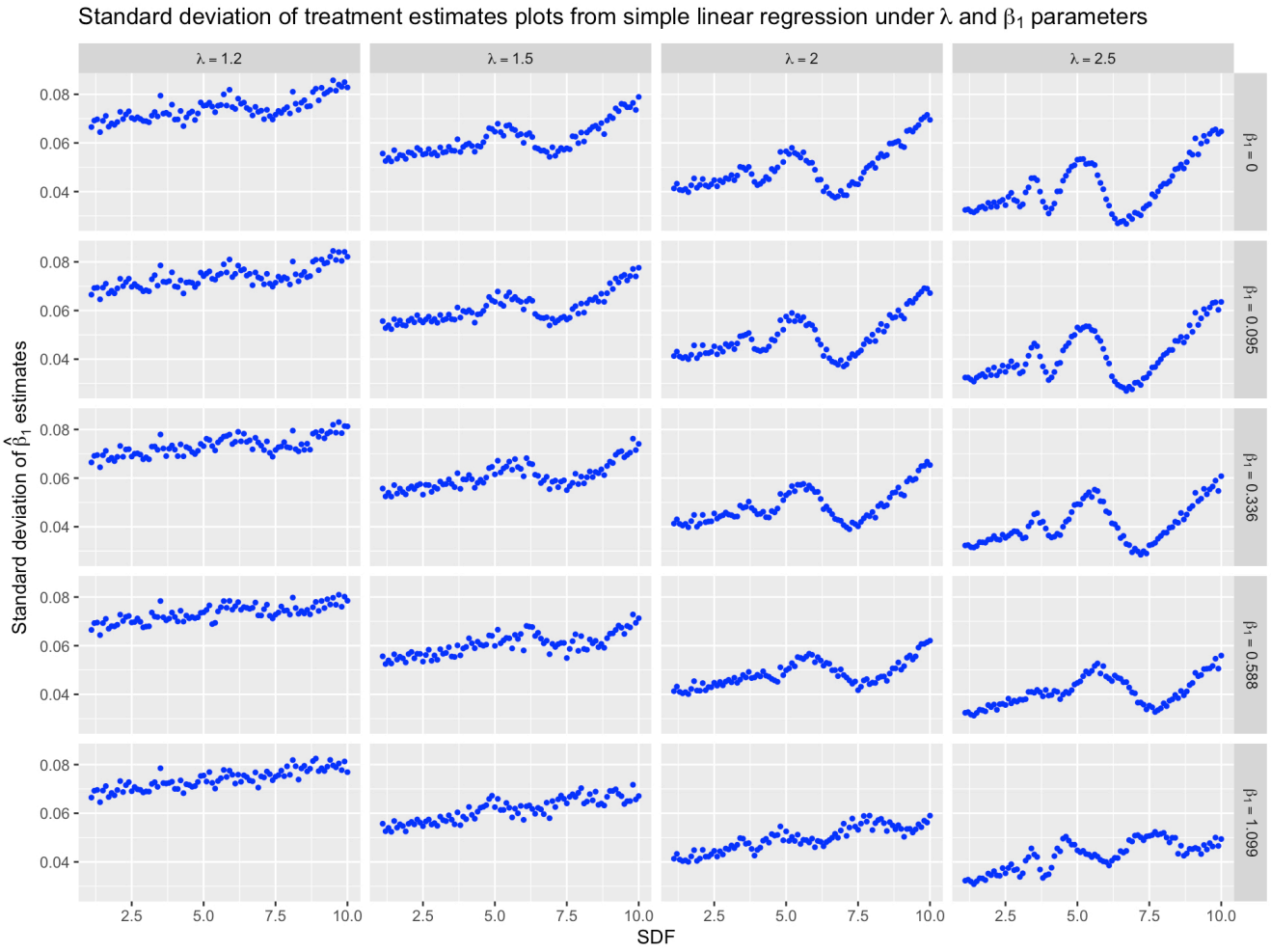

Figure 10: Scatter plots of ESDs from simple linear regression to SDFs under different λ and β_1

Depending on λ and β_1, Figure 10 shows varying relationship types between the ESD, $\sigma_e\left(\hat{\beta}_1\right)$, and the SDF. $\sigma_e\left(\hat{\beta}_1\right)$ displays a general increasing trend as the SDF increases, though the relationships are not as smooth as that of $\overline{\hat{\beta}}_1$. In comparing simulations with the same β_1, increasing λ decreases

$\sigma_e\left(\hat{\beta}_1\right)$ across all SDFs, but the periodicity becomes more pronounced towards greater SDFs. In addition, the minimum SDF in which the periodicity becomes observable also decreases.

When comparing simulations with the same λ, increasing β_1 decreases the periodicity towards linearity at smaller λ. At higher λ, the periodicity does not change in uniform fashion; the amplitude and periodic length in particular vary at different values of the SDF. However, under all combinations of λ and β_1, the apparent linear relationship between $\sigma_e\left(\hat{\beta}_1\right)$ and the SDF for all values up to 2.5 ensures that $\sigma_e\left(\hat{\beta}_1\right)$ is neither over- or underestimated in a simple linear regression utilizing this range.

iv) *Average standard errors of the log-transformed treatment estimate – ASE,* $\bar{\sigma}\left(\hat{\beta}_1\right)$

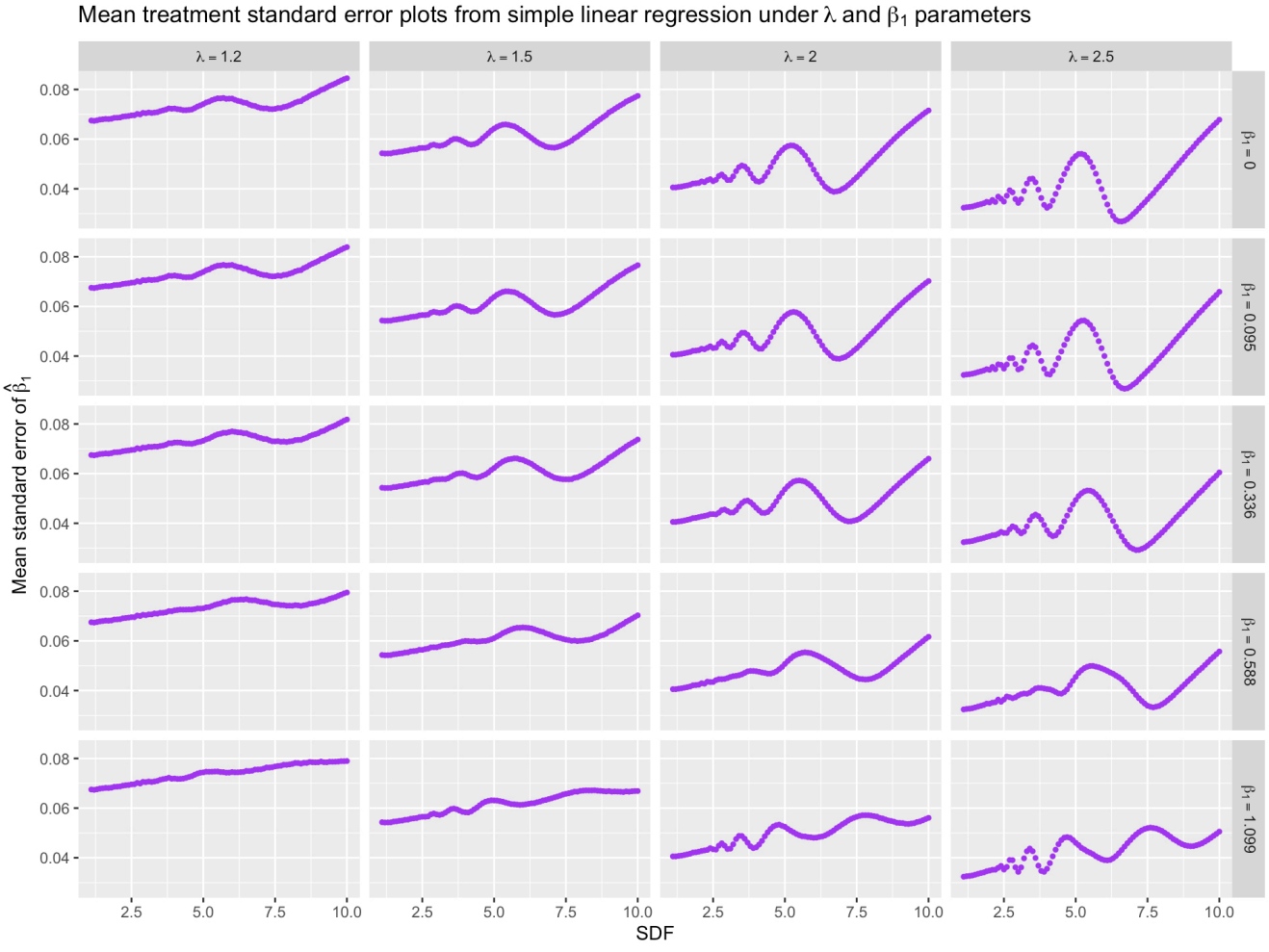

Figure 11: Scatter plots of ASEs from simple linear regression to SDFs under different λ and β_1

As with $\sigma_e\left(\hat{\beta}_1\right)$, it can be observed from Figure 11 that the ASE, $\overline{\sigma}\left(\hat{\beta}_1\right)$, follows the exact same trend with a smoother curve, $\overline{\sigma}\left(\hat{\beta}_1\right)$ being almost exactly equal to $\sigma_e\left(\hat{\beta}_1\right)$ for all settings.

3.2 **Proportional Odds Regression**

Each of the following tables displays the summarized results of the 1000 repeated trials for each proportional odds regression simulation. An unrestricted ULQ is applied to this simulation to prevent biases to the treatment parameter. Only the default SDF is used as smaller SDFs would result in too many titer categories to be analyzed, with the potential for some categories to have zero frequency.

3.2.1 [6]Unrestricted ULQ, Default SDF and No Treatment Effect

CS	TS	SS	LCO	COR	ESD	ASE	P
Low	Low	50	-0.001	0.999	0.368	0.368	0.049
Low	Low	250	-0.002	0.998	0.168	0.163	0.049
Low	Low	500	0.002	1.002	0.119	0.115	0.057
Low	High	50	-0.184	0.832	0.386	0.359	0.087
Low	High	250	-0.165	0.848	0.171	0.159	0.182
Low	High	500	-0.171	0.843	0.118	0.113	0.337
High	Low	50	0.175	1.191	0.389	0.359	0.086
High	Low	250	0.172	1.188	0.170	0.159	0.200
High	Low	500	0.167	1.182	0.124	0.113	0.335
High	High	50	-0.001	0.999	0.348	0.354	0.043
High	High	250	-0.004	0.996	0.162	0.157	0.061
High	High	500	0.003	1.003	0.111	0.111	0.052

Table 10: Simulation results under no treatment effect

[6] The full term for each statistical abbreviation in the headings of Table 10 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SS – Sample size per treatment group
LCO – Log-cumulative odds of treatment parameter
COR – Cumulative odds ratio of treatment parameter
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
P – Power of simulation study

In the case of no treatment effect, when both treatment groups have the same level of skewness, the LCO is close to 0, leading to a COR that is close to 1. However, if any treatment group has a high skewness in their distribution, this induces a non-negligible bias to the LCO. The power of the simulation also increases well beyond α, growing in severity with increasing sample sizes. Similar to the results from Table 7, both the ESD and ASE are almost equivalent to each other in all settings, and decrease with increasing sample sizes.

3.2.2 [7]Unrestricted ULQ, Default SDF and Trivial Treatment Effect

CS	TS	SS	LCO	COR	ESD	ASE	P
Low	**Low**	**50**	-0.942	0.390	0.376	0.378	0.714
Low	**Low**	**250**	-0.951	0.386	0.166	0.168	1.000
Low	**Low**	**500**	-0.949	0.387	0.119	0.119	1.000
Low	**High**	**50**	-0.680	0.507	0.391	0.366	0.462
Low	**High**	**250**	-0.661	0.516	0.169	0.163	0.978
Low	**High**	**500**	-0.660	0.517	0.115	0.115	1.000
High	**Low**	**50**	-0.487	0.614	0.387	0.362	0.280
High	**Low**	**250**	-0.485	0.615	0.169	0.161	0.833
High	**Low**	**500**	-0.484	0.616	0.121	0.113	0.986
High	**High**	**50**	-0.344	0.709	0.365	0.356	0.160
High	**High**	**250**	-0.345	0.709	0.153	0.158	0.576
High	**High**	**500**	-0.338	0.713	0.112	0.112	0.863

Table 11: Simulation results under no treatment effect

Under a trivial treatment effect, the LCO is less than 0, resulting in CORs that are less than 1. P increases with increasing sample sizes; however, it can also be observed that P becomes smaller

[7] The full term for each statistical abbreviation in the headings of Table 11 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SS – Sample size per treatment group
LCO – Log-cumulative odds of treatment parameter
COR – Cumulative odds ratio of treatment parameter
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
P – Power of simulation study

depending on the group skewness combination. The proportional odds regression results show greatly different trends for the treatment parameter than that of simple linear regression, with the treatment parameter and P showing the largest changes. Determining the effect of larger SDFs on proportional odds regression results would likewise yield different relationships between the SDF and key statistics.

3.2.3 Extended SDF Impact on Regression Statistics

The following figures in this section displays the summarized proportional odds regression results of the four key statistics outlined in Sections 2.2.1 and 2.2.2, over 500 repeated trials for each simulation to evaluate the overall impact of a larger range of SDFs on these statistics of interest, stratified by both λ and β_1 that is defined under proportional odds regression in Table 2.

i) *Log-cumulative odds of the treatment parameter – LCO,* $\overline{\hat{\beta}}_1$

Figure 12: Scatter plots of LCOs from proportional odds regression to SDFs under different λ and β_1

In Figure 12, for $\beta_1 = 0$, there are negligible deviations of $\overline{\hat{\beta}}_1$ from 0 regardless of the SDF. However, increasing λ seems to increase the spread of the LCO estimates. For simulations with the same β_1 for $\beta_1 > 0$, increasing λ increases the amplitude of the solitary periodic peak at SDFs < 2.5 which translates to a higher maximum LCO; for SDFs > 2.5, it increases the periodicity, reflecting an irregular periodic relationship with increasing negative amplitude until plateauing or becoming somewhat linear towards SDFs > 8. For simulations with the same λ for $\beta_1 > 0$, increasing β_1 increases the precision of the periodicity; this can be observed as a gradual smoothing of the points in the pattern as β_1 increases. For $\lambda < 2$, the periodicity exhibited at SDFs > 2.5 seems to gradually plateau whereas for $\lambda \geq 2$, the negative amplitude of the periodic pattern increases as evidenced by larger *y*-axis scales.

ii) *Cumulative odds ratio of the treatment parameter – COR,* $e^{\overline{\hat{\beta}}_1}$

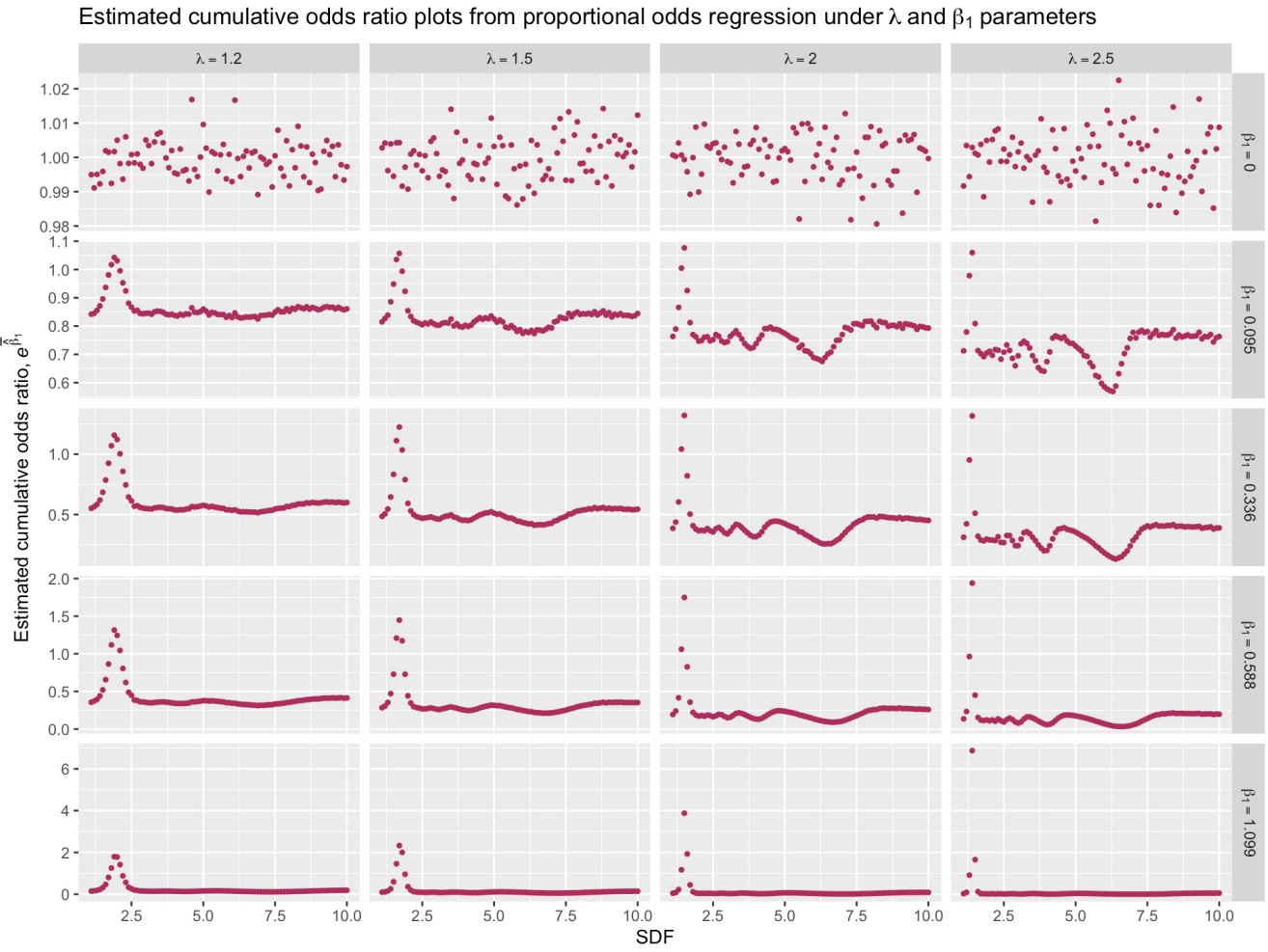

Figure 13: Scatter plots of CORs from proportional odds regression to SDFs under different λ and β_1

Like that of the relationship between the MTE and GMTR for simple linear regression, the shape and trajectory of the COR plots as shown in Figure 13 are exactly the same as that of the LCO as the COR is the exponentiation of $\overline{\hat{\beta}}_1$, but only up to a certain value of β_1. At higher SDFs, increasing β_1 further would tend $e^{\overline{\hat{\beta}}_1}$ towards 0 as the range of CORs less than 1 is limited to (0, 1). This translates to a decreasing probability of having an observed titer that is lower than any fixed observed titer level when comparing patients in the vaccine group to the control group.

The shapes and trajectories of the ESD and ASE plots for proportional odds regression mirror exactly to those of the ESD and ASE plots for simple linear regression in Figures 10 and 11 respectively. The ESD and ASE plots for proportional odds regression can be found in Appendix C for comparison with the ESD and ASE plots for simple linear regression in Figures 10 and 11.

3.3 HAI Modeling Application to Pediatric Solid Organ Transplant Data

HAI analyses are not limited to clinical trials where patients are randomized to receive either a vaccine or a placebo; they can be applied to trials that seek to evaluate immunogenicity between different vaccines for a disease, or between different dosages of a vaccine. A randomized, double-blinded Phase I clinical trial was conducted in 2014 on a group of 38 pediatric solid organ transplant (SOT) recipients between 3 and 17 years who have undergone transplant at least six months prior to their participation [1]. This study was conducted to compare, among other aims, the immunogenicity of a high-dose (HD) trivalent inactivated influenza vaccine (TIV) to that of an approved standard-dose (SD) administered to these patients. Previous research identified that SOT patients are at a higher risk of developing influenza due to medically-induced immunosuppression following SOT surgery; they also show poor immunogenic responses to the SD vaccine when compared to healthy controls [1].

An earlier Phase III study was conducted to compare the immunogenic effectiveness between recipients of the HD TIV and the SD among patients who are at least 65 years, as this age group also showed poor responses to the SD [1]. It was discovered that the HD recipients show a statistically significant higher antibody response to Influenza A antigens and increased protection against laboratory-confirmed influenza compared to those who received the SD [1]. Though the HD was subsequently licensed for use for patients in this age range, its efficacy had not been tested among pediatric SOT patients and limited data regarding influenza infections and vaccinations are in circulation for this population [1].

3.3.1 Summary of Clinical Trial Methods and Results

Patients in the SOT trial were randomized to receive either an HD TIV or the SD targeting three specific influenza strains: Influenza A H1N1, Influenza A H3N2, and Influenza B [1]. An HAI analysis was performed according to standard preparations to determine pre-vaccination and post-vaccination titers for each patient. Observed titers were calculated using the standard SDF of 2 and with a starting titer of 10. Immunogenicity was assessed through the outcomes of GMTRs, seroconversion and seroprotection, with logistic regression models used to test the effect of the TIV on seroconversion and seroprotection for each strain [1]. The main predictor of interest is the vaccine dose, and adjusting covariates include serum quantitative immunoglobulin and $CD19^+$ levels [1]. Data from 37 patients were included in the final analysis, with one patient excluded due to reception of the vaccine outside the study period [1]. No statistical baseline differences were observed for all patient demographics between the HD and SD groups [1]. From the logistic regression modeling, at $\alpha = 0.05$, the statistical analysis showed no significant differences in both post-vaccination seroconversion and seroprotection between the HD and SD treatment groups for any of the three influenza strains [1].

3.3.2 Linear and Proportional Odds Regression Modeling

It is of interest to fit the linear regression and proportional odds models to the SOT data set to determine if the HD TIV is more effective in increasing influenza antibody titers in paediatric SOT patients compared to the SD, and if similar results to that obtained from logistic regressions used for seroconversion and seroprotection could be concluded. As demonstrated through the simulation study that the default SDF of 2 does not heavily impact the treatment effect of the regression model nor the standard error of the predictor compared to smaller SDFs, both models can be utilized with this data set without sacrificing efficiency. As the TIV study was a Phase I clinical trial, the study would only have a small sample size of paediatric SOT patients. With data from 37 patients, it is imperative that variable selection should be done carefully so as to not overfit either regression model to the data. Harrell described a method of calculating the number of predictors to be included in a regression model using the concept of a "limiting sample size," denoted as m [2]. To determine m, the frequency distribution of the outcome variable – the post-vaccination observed HAI titer in the case of the SOT data – must first be evaluated. Table 12 displays the frequencies of observed titer levels of all 37 patients in the data set stratified by influenza strain.

Influenza strain	Observed serum HAI titers											Total
	5	10	20	40	80	160	320	640	1280	2560	5120	
A H1N1	1	0	2	1	1	0	8	8	7	4	5	37
A H3N2	0	1	6	3	7	6	4	7	3	0	0	37
B	1	10	5	10	4	3	1	3	0	0	0	37

Table 12: Frequency distribution of strain-specific post-vaccination observed titers of SOT data

Using the frequency distribution, *m* can then be calculated based on the data type of the outcome variable. Harrell summarized the criteria and calculation procedure for *m* in Table 13 below [2]. Following this, the number of predictors to be included in each regression model to be used can thus be specified; from previous regression modeling studies, Harrell proposed that fitted regression models are likely to be reliable when the total number of predictors is less than $\frac{m}{15}$, otherwise known as the "15:1 rule of thumb." [1] With these guidelines, *m* is obtained for the SOT data for each influenza strain using the formulas provided in Table 13 and the frequency distribution of the patient post-vaccination observed HAI titer in Table 12. The total number of predictors to be included in the regression model is then calculated using the rule of thumb above.

Regression model	Outcome variable	Limiting sample size calculation	Influenza strain	Limiting sample size, *m*	Total predictors
Linear	Continuous	n	All	37	2
Proportional odds	Ordinal with *K* categories	$n - \frac{1}{n^2}\sum_{i=1}^{K} n_i^3$	A H1N1	35.9	2
			A H3N2	36.1	2
			B	35.4	2

Table 13: Limiting sample size calculations for regression modeling variable selection [2]

For both the linear and proportional odds regression models that will be fitted to the data, a maximum of two predictors are to be used. In addition to the vaccine dose – the main predictor of interest – the additional adjusting covariate to be included in the regression models is the log-transformed pre-vaccination observed HAI titer. This is because each child in the SOT data set starts

with a different pre-vaccination titer, and not adjusting for this covariate would result in biased estimates and possibly spurious associations. While gender and race could also be adjusting covariates, models with the the pre-vaccination titer fit the data better compared to models with these variables. Hence, the models fitted to the data are:

Linear regression: $E\left[\log Y \mid X = x, Z = z\right] = \beta_0^\star + \beta_1^\star x + \beta_2^\star \log z$ (11)

Proportional odds regression: $\log\left[\dfrac{P(Y \le y)}{1 - P(Y \le y)} \,\middle|\, X = x, Z = z\right] = \beta_0^* + \beta_1^* x + \beta_2^* \log z$ (12)

where in models (11) and (12) above:

- Y is the post-vaccination observed serum HAI titer
- $P(Y \le y)$ is the probability that the post-vaccination titer is at most the observed titer level y
- X is the vaccine dose level
- Z is the pre-vaccination observed serum HAI titer
- $\beta_0^\star$ is the mean log-transformed post-vaccination titer for patients in the SD group with a pre-vaccination titer of 1
- $\beta_1^\star$ is the difference in the mean log-transformed post-vaccination titer when comparing patients in the HD group to the SD group with the same pre-vaccination titer
- $\beta_2^\star \log(1 + q)$ is the difference in the mean log-transformed post-vaccination titer when comparing patients in the the same treatment group that differ in their pre-vaccination titer by $(100 \times q)\%$
- β_0^* is the log-cumulative odds of being in a post-vaccination titer level of at most y for patients in the SD group with a pre-vaccination titer of 1
- β_1^* is the difference in the log-cumulative odds of being in a post-vaccination titer level of at most y when comparing patients in the HD group to the SD group with the same pre-vaccination titer
- $\beta_2^* \log(1 + q)$ is the difference in the log-cumulative odds of being in a post-vaccination titer level of at most y when comparing patients in the the same treatment group that differ in their pre-vaccination titer by $(100 \times q)\%$

The tables in the following pages summarize the linear regression results for each influenza strain, with the GMT or GMTR included for each parameter alongside a corresponding 95% confidence interval (CI). The vaccine dose predictor is highlighted in yellow in each table.

3.3.3 Regression Modeling Results

i) [8]*Linear regression modeling*

Parameter (Influenza A H1N1)	MTE (SE)	GMT/GMTR [95% CI]	*p*-value
Intercept	4.395 (0.657)	81.00 [21.32, 307.81]	< 0.001
Dose = HD (Reference: SD)	0.989 (0.488)	2.688 [0.997, 7.244]	0.0506
Pre-vaccination log titer	0.335 (0.132)	1.398 [1.070, 1.826]	0.0156

Table 14: Linear regression results for testing TIV immunogenicity on Influenza A H1N1

Parameter (Influenza A H3N2)	MTE (SE)	GMT/GMTR [95% CI]	*p*-value
Intercept	1.937 (0.673)	6.941 [1.767, 27.26]	0.00687
Dose = HD (Reference: SD)	0.643 (0.387)	1.901 [0.866, 4.173]	0.106
Pre-vaccination log titer	0.656 (0.138)	1.928 [1.458, 2.549]	< 0.001

Table 15: Linear regression results for testing TIV immunogenicity on Influenza A H3N2

Parameter (Influenza B)	MTE (SE)	GMT/GMTR [95% CI]	*p*-value
Intercept	0.593 (0.414)	1.810 [0.780, 4.200]	0.161
Dose = HD (Reference: SD)	0.109 (0.262)	1.115 [0.654, 1.900]	0.682
Pre-vaccination log titer	0.987 (0.123)	2.683 [2.091, 3.442]	< 0.001

Table 16: Linear regression results for testing TIV immunogenicity on Influenza B

From the linear regression model fitted to the SOT data for each influenza strain, the GMTR in comparing patients in the HD group to those in the SD group – adjusted for the log-transformed pre-vaccination observed titer – is higher than 1. This means that the GMT of patients in the HD group is higher than that of the SD group by the following:

[8] The full term for each statistical abbreviation in the headings of Tables 14, 15, and 16 are as follows:
MTE – Mean log-transformed treatment estimate
SE – Standard error of MTE
GMT – Geometric mean titer
GMTR – Geometric mean titer ratio

- 2.69 times higher for Influenza A H1N1
- 1.9 times higher for Influenza A H3N2
- 1.12 times higher for Influenza B

The *p*-value of the vaccine dose parameter, however, is higher than 0.05 for each strain. Thus, there is insufficient evidence to suggest that the HD TIV is more effective in raising antibody titers in pediatric SOT patients than the SD for the three influenza strains. This conclusion can also be explained by the fact that the CI obtained for the treatment parameter in each regression model does not rule out a GMTR of 1.

The following tables summarize the proportional odds regression results for each influenza strain, with the cumulative odds (CO) or COR included for each parameter alongside a corresponding 95% CI. The predictor of interest is highlighted in yellow in each table.

ii) [9]*Proportional odds regression modeling*

Parameter (Influenza A H1N1)	LCO (SE)	CO/COR [95% CI]	*p*-value
Intercept: *Post-titer* ≤ 5	-2.816 (1.069)	0.0599 [0.00737, 0.486]	NA
Intercept: *Post-titer* ≤ 20	-1.592 (0.684)	0.204 [0.0533, 0.778]	0.0199
Intercept: *Post-titer* ≤ 40	-1.248 (0.624)	0.287 [0.0844, 0.976]	0.0457
Intercept: *Post-titer* ≤ 80	-0.982 (0.590)	0.375 [0.118, 1.191]	0.0962
Intercept: *Post-titer* ≤ 320	0.352 (0.536)	1.422 [0.497, 4.071]	0.511
Intercept: *Post-titer* ≤ 640	1.336 (0.576)	3.804 [1.229, 11.77]	0.0205
Intercept: *Post-titer* ≤ 1280	2.300 (0.652)	9.975 [2.781, 35.78]	< 0.001
Intercept: *Post-titer* ≤ 2560	3.055 (0.739)	21.21 [4.984, 90.28]	< 0.001
Dose = HD (Reference: SD)	-1.142 (0.618)	0.319 [0.0951, 1.071]	0.0646
Pre-vaccination log titer	-0.00134 (0.00109)	0.999 [0.997, 1.001]	0.218

Table 17: Proportional odds regression results for testing TIV immunogenicity on Influenza A H1N1

[9] The full term for each statistical abbreviation in the headings of Tables 17, 18, and 19 are as follows:
LCO – Log-cumulative odds of treatment parameter
SE – Standard error of LCO
CO – Cumulative odds of treatment parameter
COR – Cumulative odds ratio of treatment parameter

Parameter (Influenza A H3N2)	LCO (SE)	CO/COR [95% CI]	*p*-value
Intercept: *Post-titer* ≤ 10	-2.490 (1.136)	0.0829 [0.00895, 0.768]	NA
Intercept: *Post-titer* ≤ 20	-0.183 (0.629)	0.833 [0.243, 2.856]	0.771
Intercept: *Post-titer* ≤ 40	0.387 (0.614)	1.473 [0.442, 4.911]	0.528
Intercept: *Post-titer* ≤ 80	1.400 (0.643)	4.056 [1.151, 14.30]	0.0294
Intercept: *Post-titer* ≤ 160	2.272 (0.701)	9.697 [2.454, 38.32]	0.00119
Intercept: *Post-titer* ≤ 320	2.987 (0.771)	19.83 [4.374, 89.93]	< 0.001
Intercept: *Post-titer* ≤ 640	4.920 (1.111)	137.02 [15.53, 1209.03]	< 0.001
Dose = HD (Reference: SD)	-1.123 (0.649)	0.325 [0.0912, 1.160]	0.0834
Pre-vaccination log titer	-0.00811 (0.00243)	0.992 [0.987, 0.997]	< 0.001

Table 18: Proportional odds regression results for testing TIV immunogenicity on Influenza A H3N2

Parameter (Influenza B)	LCO (SE)	CO/COR [95% CI]	*p*-value
Intercept: *Post-titer* ≤ 5	-2.217 (1.083)	0.109 [0.0130, 0.910]	0.0407
Intercept: *Post-titer* ≤ 10	0.780 (0.597)	2.182 [0.677, 7.036]	0.191
Intercept: *Post-titer* ≤ 20	1.566 (0.625)	4.789 [1.407, 16.30]	0.0122
Intercept: *Post-titer* ≤ 40	3.280 (0.803)	26.57 [5.509, 128.11]	< 0.001
Intercept: *Post-titer* ≤ 80	4.296 (0.984)	73.37 [10.66, 504.83]	< 0.001
Intercept: *Post-titer* ≤ 160	5.779 (1.352)	323.38 [22.86, 4573.65]	< 0.001
Intercept: *Post-titer* ≤ 320	6.837 (1.607)	931.46 [39.92, 21734]	< 0.001
Dose = HD (Reference: SD)	-0.433 (0.678)	0.649 [0.172, 2.450]	0.523
Pre-vaccination log titer	-0.0692 (0.0202)	0.933 [0.897, 0.971]	< 0.001

Table 19: Proportional odds regression results for testing TIV immunogenicity on Influenza B

From the proportional odds regression model fitted to the SOT data for each influenza strain, the COR of having an observed post-vaccination titer that is lower than any fixed observed titer level when comparing patients in the HD group to those in the SD group – adjusted for the log-transformed pre-vaccination observed titer – is lower than 1. This means that the CO of having an observed post-

vaccination titer that is lower than any fixed observed titer level for patients in the HD group is lower than that of the SD group by the following:

- 68.9% lower for Influenza A H1N1
- 67.5% lower for Influenza A H3N2
- 35.1% lower for Influenza B

The *p*-value of the vaccine dose parameter, however, is higher than 5% for each strain. Thus, there is insufficient evidence to suggest that the HD TIV is more effective in raising antibody titers in pediatric SOT patients than the SD for the three influenza strains. This conclusion can also be explained by the fact that the CI obtained for the treatment parameter in each regression model does not rule out a COR of 1.

Comparing the conclusions from these regression models to those obtained from the logistic regression models for seroconversion and seroprotection, these results could be said to be consistent as a non-significant result for increased antibody titers for the HD TIV would highly correspond to non-significant results for increased seroconversion and seroprotection with the same dose.

CHAPTER 4

DISCUSSION

At first glance, it would seem counterintuitive that decreasing the SDF in order to obtain more granular observed titer levels does little to improve the already small bias in the MTE for linear regression when there is any level of treatment effect. The extended SDF impact simulation in Section 3.1.4 helps to explain numerically this odd occurrence. As the SDF decreases from 10 up to 1.1, a damping effect is applied on the periodic relationship between the SDF and the MTE; towards SDFs below 2.5, the periodicity disappears almost to the point of constancy. This implies that much finer SDFs do not seem to significantly improve the MTE, which practically translates to an unnecessary increase in the amount of serum dilutions needed for each distinct observed titer level. However, this does not change the fact that higher SDFs can either drastically over- or underestimate the TTP depending on how large of an SDF is pre-selected compared to the default SDF of 2 that is commonly used in HAI analysis.

When a restricted MTD is imposed on the HAI procedure, the MTE bias does increase slightly at much smaller SDFs, but its impact on the reduction of CP is considerable. Using a low ULQ in an HAI analysis is ill-advised as it would severely reduce the ability to obtain MTEs that are close to the TTP using observed titers. It would seem that low ULQ has a much higher impact than smaller SDFs in inducing bias to the estimates of regression modeling.

Finer SDFs also do not significantly reduce the ESD or ASE compared to the default SDF. On the other hand, increasing group sample sizes does decrease both the ESD and ASE across all simulation settings for both regression types. This is in accord with the Law of Large Numbers which states that the estimated average of a statistic becomes closer to its expectation as the number of trials, or samples, increases. Increasing the treatment group sample sizes in a clinical trial is thus encouraged so as to achieve more consistent results to the underlying true parameters. However, when concerning a Phase I clinical trial, sample sizes are a massive limiting factor due to the nature and aim of this phase that is more concerned with patient safety, in addition to the per-patient resources and costs that must be expended. As with the case of the TIV study, only 38 patients in total were enrolled in the clinical trial; with such a small sample size, it would be difficult to truly gauge the true immunogenic effect

from HAI regression modeling. Despite this limitation, it is encouraging to note that the simulation results at the smallest per-treatment group sample size yield similar MTEs and CP when compared to larger sample sizes.

Concerning proportional odds regression, the results are heavily influenced by treatment group skewness compared to linear regression. This is largely based on the fact that proportional odds regression accounts for the ordinal nature of the data and not their actual values. In this study, based on the set parameters for skewness in the Weibull data generating mechanism, simulating data for any treatment group with high skewness would inevitably lead to more total data with lower titers compared to that with low skewness. Nevertheless, a proportional odds model is still highly viable in the HAI analysis as it answers different statistical questions to that of linear regression.

With many variations among laboratories and research centers in determining the initial HAI settings such as the starting dilution titer, SDF, and ULQ before the HAI is conducted, it is inevitable that complications would arise in verifying and reproducing the results between different entities. Zacour et al. reported that "poor reproducibility of HAI results from one laboratory to another is widely cited, limiting comparisons between candidate vaccines in different clinical trials and posing challenges for licensing authorities." [15] It is imperative that HAI standardization efforts be streamlined in order to improve inter-laboratory reproducibility and communication, especially in the event of a pandemic that would demand timely and accurate research to better facilitate testing and approval of medical interventions [15]. Good standardizations can serve to greatly improve future immunogenicity analyses that can lead to better informed decision-making with regard to vaccine licensure.

CHAPTER 5

CONCLUSION

Regression modeling is a viable alternative to hypothesis testing in evaluating vaccine immunogenicity using an HAI analysis. In linear regression modeling, both the bias and the standard error of the treatment effect – the main predictor of interest in HAI analysis – have a periodic relationship to the SDF used in discretizing serum titers. This periodic relationship, however, is not prominent for SDFs up to 2.5, leading to negligible changes to the bias and standard error of the treatment effect. On the other hand, for proportional odds regression modeling, the periodic relationship is more extreme at small SDFs, though the relationships of both the standard deviation and standard error of main predictor estimates to the SDF are exactly the same with that of linear regression. With current HAI analytical practice that typically uses an SDF of 2 in preparing blood serum samples, there is almost no risk of loss of information when using observed titer values as compared to using more granulated titer categories with smaller SDFs. However, standardizing the SDF selection before the HAI analysis is conducted, especially in influenza serology, and allowing for an unrestricted ULQ would lead to greater reproducibility in results, reduction in unnecessary information loss, and better collaborative efforts in licensing novel or improved vaccines.

REFERENCES

[1] GiaQuinta S, Michaels MG, McCullers JA, et al. Randomized, Double-blind Comparison of Standard-dose vs. High-dose Trivalent Inactivated Influenza Vaccine in Pediatric Solid Organ Transplant Patients. *Pediatr Transplant*. 2015;19(2): 219-228. doi:10.1111/petr.12419

[2] Harrell FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*. 2nd ed. Springer International Publishing; 2015.

[3] Hirst GK. The Quantitative Determination of Influenza Virus and Antibodies by Means of Red Cell Agglutination. *J Exp Med*. 1942;75(1): 49-64. doi: 10.1084/jem.75.1.49

[4] Hsu JL. A Brief History of Vaccines: Smallpox to the Present. *S D Med*. 2013;Spec no: 33-37.

[5] Kaufmann L, Syedbasha M, Vogt D, et al. An Optimized Hemagglutination Inhibition (HI) Assay to Quantify Influenza-specific Antibody Titers. *J Vis Exp*. 2017;(130): 55833. Published 2017 Dec 1. doi: 10.3791/55833

[6] Kay AW, Blish CA. Immunogenicity and Clinical Efficacy of Influenza Vaccination in Pregnancy. *Front Immunol*. 2015;6: 289. Published 2015 Jun 4. doi: 10.3389/ fimmu.2015.00289

[7] Liu Q, Shepherd BE, Li C, Harrell FE Jr. Modeling Continuous Response Variables Using Ordinal Regression. *Stat Med*. 2017;36(27): 4316-4335. doi: 10.1002/sim.7433

[8] Morabia A. Edward Jenner's 1798 Report of Challenge Experiments Demonstrating the Protective Effects of Cowpox Against Smallpox. *J R Soc Med*. 2018;111(7): 255-257. doi: 10.1177/0141076818783658

[9] Nguyen M, Fries K, Khoury R, et al. Automated Imaging and Analysis of the Hemagglutination Inhibition Assay. *J Lab Autom*. 2016;21(2): 287-296. doi: 10.1177/2211068215610061

[10] Nhat NTD, Todd S, de Bruin E, et al. Structure of General-population Antibody Titer Distributions to Influenza A Virus. *Sci Rep*. 2017;7(1): 6060. Published 2017 Jul 20. doi: 10.1038/s41598-017-06177-0

[11] Noah DL, Hill H, Hines D, White EL, Wolff MC. Qualification of the Hemagglutination Inhibition Assay in Support of Pandemic Influenza Vaccine Licensure. *Clin Vaccine Immunol*. 2009;16(4): 558-566. doi: 10.1128/CVI.00368-08

[12] Pulendran B, Ahmed R. Immunological Mechanisms of Vaccination. *Nat Immunol*. 2011;12(6): 509-517. doi: 10.1038/ni.2039

[13] Rencher AC, Schaalje GB. *Linear Models in Statistics*. 2nd ed. John Wiley & Sons, Inc.; 2008.

[14] Singh K, Mehta S. The Clinical Development Process for a Novel Preventive Vaccine: An Overview. *J Postgrad Med*. 2016;62(1): 4-11. doi: 10.4103/0022-3859.173187

[15] Zacour M, Ward BJ, Brewer A, Tang P, Boivin G, Li Y, Warhuus M, McNeil SA, LeBlanc JJ, Hatchette TF. Standardization of Hemagglutination Inhibition Assay for Influenza Serology Allows for High Reproducibility Between Laboratories. *Clinical and Vaccine Immunology*. 2016; 23(3): 236-242. doi: 10.1128/cvi.00613-15

APPENDICES

Appendix A: Simple Linear Regression with Restricted ULQ and Default SDF

[10]*No treatment effect*

CS	TS	SS	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	Low	50	0.001	0.001	0.147	0.144	0.947	0.053	0.001	0.153	0.150	0.946	0.053	0.927
Low	Low	250	-0.000	0.001	0.067	0.065	0.943	0.057	0.001	0.069	0.068	0.946	0.055	0.938
Low	Low	500	-0.001	-0.001	0.047	0.046	0.946	0.054	-0.001	0.048	0.048	0.947	0.054	0.935
Low	High	50	-0.009	-0.006	0.211	0.210	0.947	0.054	0.005	0.212	0.210	0.944	0.055	0.994
Low	High	250	-0.006	-0.008	0.096	0.095	0.947	0.054	0.003	0.096	0.095	0.942	0.055	1.007
Low	High	500	-0.005	-0.005	0.066	0.067	0.951	0.050	0.007	0.066	0.067	0.947	0.050	1.002
High	Low	50	0.004	0.010	0.217	0.210	0.939	0.060	-0.001	0.217	0.209	0.936	0.065	1.004
High	Low	250	0.008	0.006	0.096	0.095	0.947	0.053	-0.005	0.095	0.095	0.944	0.052	1.013
High	Low	500	0.007	0.006	0.067	0.067	0.947	0.050	-0.005	0.067	0.067	0.948	0.050	1.015
High	High	50	0.001	-0.002	0.264	0.260	0.945	0.055	-0.003	0.260	0.256	0.944	0.056	1.034
High	High	250	0.000	-0.001	0.120	0.118	0.949	0.051	-0.001	0.118	0.116	0.950	0.050	1.035
High	High	500	0.001	0.001	0.084	0.083	0.947	0.054	0.000	0.083	0.082	0.946	0.054	1.032

Table 20: Simulation results under no treatment effect

Under no treatment effect, the regression models utilizing both latent and observed titers perform almost equally to each other, with barely noticeable differences across all statistical measurements and under different group skewness combinations. The MTE across all trials is also

[10] The full term for each statistical abbreviation in the headings of Table 20 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SS – Sample size per treatment group
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

fairly close to the TTP of 0, corresponding to GMTRs that are very close to 1. The regression model with latent titers also seems to be more efficient when both treatment groups have low skewness, whereas the RE shifts slightly in favor of the regression model with observed titers when either group has a highly skewed distribution. The ESD and ASE are almost equal to each other, and both decrease when the group sample sizes increase across all group skewness combinations. At $\alpha = 0.05$, the CP and P are fairly close to the expected proportions of 0.95 and 0.05 respectively.

[11]*Moderate treatment effect*

CS	TS	SS	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**50**	1.351	1.352	0.146	0.146	0.945	1.000	1.350	0.150	0.151	0.949	1.000	0.951
Low	**Low**	**250**	1.351	1.351	0.066	0.066	0.952	1.000	1.349	0.069	0.068	0.949	1.000	0.941
Low	**Low**	**500**	1.352	1.350	0.047	0.047	0.948	1.000	1.348	0.049	0.048	0.945	1.000	0.923
Low	**High**	**50**	1.094	1.093	0.192	0.189	0.946	0.999	1.071	0.192	0.188	0.945	0.999	0.995
Low	**High**	**250**	1.097	1.095	0.086	0.085	0.950	1.000	1.073	0.086	0.085	0.940	1.000	1.013
Low	**High**	**500**	1.094	1.094	0.061	0.060	0.949	1.000	1.073	0.061	0.060	0.936	1.000	1.005
High	**Low**	**50**	1.357	1.359	0.214	0.211	0.943	1.000	1.345	0.213	0.211	0.941	1.000	1.006
High	**Low**	**250**	1.357	1.357	0.095	0.096	0.952	1.000	1.343	0.095	0.095	0.951	1.000	1.009
High	**Low**	**500**	1.358	1.358	0.067	0.068	0.951	1.000	1.344	0.067	0.067	0.946	1.000	1.000
High	**High**	**50**	1.103	1.101	0.250	0.243	0.941	0.993	1.068	0.245	0.239	0.941	0.991	1.041
High	**High**	**250**	1.103	1.103	0.109	0.110	0.955	1.000	1.069	0.108	0.108	0.943	1.000	1.032
High	**High**	**500**	1.101	1.102	0.078	0.078	0.951	1.000	1.069	0.077	0.076	0.931	1.000	1.036

Table 21: Simulation results under a moderate treatment effect

[11] The full term for each statistical abbreviation in the headings of Table 21 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SS – Sample size per treatment group
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

Under a moderate treatment effect, as is the case with no treatment effect, the regression models utilizing both latent and observed titers also perform almost equally to each other, though the magnitude of the bias in MTE for the model with observed titers compared to that with latent values is slightly greater when either or both treatment groups have a higher skewness. Similar to the case of no treatment effect, the regression model with latent titers seems to be more efficient when both treatment groups have low skewness, whereas the RE shifts only slightly in favor of the regression model with observed titers when either group has a highly skewed distribution. Both the ESD and ASE are almost equal to each other for both latent and observed titers as well as across titer types. Both also become smaller when the group sample sizes increase. In most trials, though the CP remains fairly close to 0.95, P jumps significantly towards 1; this may be due to the simulated MTEs with corresponding estimated GMTRs between $e^{1.094} \approx 3$ and $e^{1.358} \approx 3.9$ from the results.

Appendix B: Simple Linear Regression with Restricted ULQ, Large Treatment Effect, and Decreasing SDF

[12]*Small sample sizes* (*50 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	1.718	1.715	0.147	0.146	0.947	1.000	1.685	0.149	0.148	0.939	1.000	0.979
Low	**Low**	**1.8**	1.717	1.721	0.149	0.146	0.945	1.000	1.715	0.151	0.149	0.945	1.000	0.965
Low	**Low**	**1.6**	1.716	1.715	0.146	0.146	0.948	1.000	1.676	0.144	0.144	0.936	1.000	1.033
Low	**Low**	**1.4**	1.719	1.718	0.147	0.146	0.947	1.000	1.679	0.143	0.143	0.940	1.000	1.056
Low	**Low**	**1.2**	1.719	1.715	0.148	0.146	0.949	1.000	1.642	0.141	0.139	0.914	1.000	1.102
Low	**High**	**2.0**	1.579	1.582	0.175	0.170	0.941	1.000	1.525	0.171	0.166	0.937	1.000	1.046
Low	**High**	**1.8**	1.579	1.582	0.173	0.170	0.945	1.000	1.561	0.173	0.170	0.944	1.000	1.001
Low	**High**	**1.6**	1.580	1.575	0.170	0.170	0.950	1.000	1.510	0.163	0.163	0.937	1.000	1.088
Low	**High**	**1.4**	1.580	1.587	0.174	0.170	0.946	1.000	1.521	0.166	0.163	0.934	1.000	1.091
Low	**High**	**1.2**	1.580	1.578	0.175	0.170	0.938	1.000	1.478	0.163	0.159	0.901	1.000	1.147
High	**Low**	**2.0**	1.721	1.724	0.214	0.212	0.945	1.000	1.683	0.211	0.208	0.940	1.000	1.031
High	**Low**	**1.8**	1.724	1.726	0.216	0.212	0.944	1.000	1.708	0.214	0.210	0.943	1.000	1.016
High	**Low**	**1.6**	1.724	1.721	0.215	0.212	0.947	1.000	1.672	0.209	0.206	0.937	1.000	1.054
High	**Low**	**1.4**	1.723	1.727	0.216	0.212	0.942	1.000	1.679	0.210	0.206	0.934	1.000	1.051
High	**Low**	**1.2**	1.721	1.719	0.215	0.211	0.943	1.000	1.642	0.206	0.203	0.921	1.000	1.087
High	**High**	**2.0**	1.588	1.587	0.231	0.229	0.945	1.000	1.517	0.223	0.222	0.928	1.000	1.065
High	**High**	**1.8**	1.587	1.587	0.233	0.228	0.943	1.000	1.555	0.228	0.225	0.942	1.000	1.043
High	**High**	**1.6**	1.588	1.587	0.235	0.229	0.944	1.000	1.512	0.227	0.221	0.925	1.000	1.077
High	**High**	**1.4**	1.589	1.583	0.228	0.229	0.954	1.000	1.511	0.219	0.220	0.931	1.000	1.081
High	**High**	**1.2**	1.585	1.591	0.228	0.229	0.949	1.000	1.486	0.216	0.217	0.917	1.000	1.115

Table 22: Simulation results under small sample sizes

[12] The full term for each statistical abbreviation in the headings of Table 22 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

[13]*Medium sample sizes* (*250 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	1.719	1.719	0.065	0.066	0.955	1.000	1.689	0.065	0.067	0.933	1.000	0.988
Low	**Low**	**1.8**	1.717	1.718	0.066	0.066	0.949	1.000	1.712	0.067	0.067	0.948	1.000	0.970
Low	**Low**	**1.6**	1.720	1.716	0.067	0.066	0.945	1.000	1.677	0.066	0.065	0.901	1.000	1.041
Low	**Low**	**1.4**	1.717	1.720	0.067	0.066	0.946	1.000	1.680	0.065	0.065	0.912	1.000	1.047
Low	**Low**	**1.2**	1.717	1.717	0.067	0.066	0.946	1.000	1.645	0.064	0.063	0.785	1.000	1.107
Low	**High**	**2.0**	1.581	1.581	0.078	0.077	0.942	1.000	1.524	0.076	0.075	0.887	1.000	1.063
Low	**High**	**1.8**	1.581	1.580	0.077	0.077	0.950	1.000	1.559	0.077	0.077	0.941	1.000	1.005
Low	**High**	**1.6**	1.581	1.582	0.078	0.077	0.943	1.000	1.517	0.075	0.074	0.868	1.000	1.080
Low	**High**	**1.4**	1.580	1.578	0.077	0.077	0.951	1.000	1.514	0.074	0.073	0.856	1.000	1.090
Low	**High**	**1.2**	1.580	1.580	0.077	0.077	0.951	1.000	1.481	0.071	0.072	0.730	1.000	1.163
High	**Low**	**2.0**	1.721	1.723	0.097	0.096	0.946	1.000	1.680	0.096	0.094	0.925	1.000	1.038
High	**Low**	**1.8**	1.722	1.724	0.097	0.096	0.947	1.000	1.707	0.096	0.095	0.940	1.000	1.023
High	**Low**	**1.6**	1.723	1.726	0.097	0.096	0.949	1.000	1.678	0.094	0.093	0.917	1.000	1.049
High	**Low**	**1.4**	1.727	1.721	0.096	0.096	0.950	1.000	1.674	0.093	0.093	0.904	1.000	1.059
High	**Low**	**1.2**	1.723	1.724	0.096	0.096	0.950	1.000	1.647	0.092	0.092	0.863	1.000	1.080
High	**High**	**2.0**	1.587	1.586	0.103	0.103	0.953	1.000	1.517	0.100	0.100	0.891	1.000	1.063
High	**High**	**1.8**	1.586	1.587	0.104	0.103	0.949	1.000	1.555	0.101	0.102	0.939	1.000	1.041
High	**High**	**1.6**	1.581	1.587	0.103	0.104	0.950	1.000	1.513	0.099	0.100	0.895	1.000	1.084
High	**High**	**1.4**	1.585	1.588	0.103	0.103	0.950	1.000	1.515	0.099	0.099	0.892	1.000	1.077
High	**High**	**1.2**	1.587	1.589	0.106	0.103	0.943	1.000	1.484	0.101	0.098	0.808	1.000	1.114

Table 23: Simulation results under medium sample sizes

[13] The full term for each statistical abbreviation in the headings of Table 23 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

In this simulation setting, from Table 22, there is a small but noticeable bias in MTE in all trials for the model with observed titers compared to latent titers; the level of bias depends on the group skewness combinations. The ESD and the ASE are still equivalent to each other, and also show increasing trends according to the order of group skewness combination shown above. The CP starts to slowly decrease from 0.95 at smaller SDFs. Conversely, towards smaller SDFs, the RE shows an inconsistent increasing trend in favor of the model with observed titers.

Table 23 shows similar trends towards smaller SDFs for the RE and bias of the MTE. Both the ESD and ASE are still relatively equal to each other, but are much smaller at medium sample sizes. The CP shows an accelerated decreasing rate from the expected probability of 0.95 at much lower SDFs.

Appendix C: Simple Linear Regression with Unrestricted ULQ and Decreasing SDF

[14]*Large treatment effect and small sample sizes* (*50 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	1.718	1.716	0.148	0.146	0.944	1.000	1.714	0.153	0.151	0.947	1.000	0.939
Low	**Low**	**1.8**	1.717	1.717	0.149	0.146	0.942	1.000	1.716	0.153	0.150	0.944	1.000	0.946
Low	**Low**	**1.6**	1.718	1.718	0.147	0.146	0.946	1.000	1.718	0.150	0.148	0.946	1.000	0.968
Low	**Low**	**1.4**	1.718	1.717	0.148	0.146	0.943	1.000	1.716	0.149	0.147	0.943	1.000	0.982
Low	**Low**	**1.2**	1.718	1.718	0.150	0.146	0.942	1.000	1.718	0.151	0.147	0.941	1.000	0.998
Low	**High**	**2.0**	1.578	1.581	0.173	0.170	0.946	1.000	1.580	0.177	0.174	0.947	1.000	0.952
Low	**High**	**1.8**	1.580	1.582	0.173	0.170	0.947	1.000	1.582	0.176	0.173	0.947	1.000	0.967
Low	**High**	**1.6**	1.582	1.576	0.174	0.170	0.941	1.000	1.576	0.176	0.172	0.938	1.000	0.979
Low	**High**	**1.4**	1.580	1.580	0.171	0.170	0.946	1.000	1.579	0.171	0.171	0.947	1.000	0.991
Low	**High**	**1.2**	1.581	1.579	0.173	0.170	0.944	1.000	1.579	0.173	0.170	0.942	1.000	1.000
High	**Low**	**2.0**	1.725	1.726	0.219	0.212	0.942	1.000	1.713	0.218	0.211	0.942	1.000	1.010
High	**Low**	**1.8**	1.722	1.722	0.217	0.212	0.941	1.000	1.710	0.215	0.211	0.940	1.000	1.015
High	**Low**	**1.6**	1.726	1.716	0.210	0.211	0.944	1.000	1.707	0.209	0.210	0.945	1.000	1.013
High	**Low**	**1.4**	1.724	1.724	0.214	0.212	0.945	1.000	1.715	0.212	0.210	0.944	1.000	1.025
High	**Low**	**1.2**	1.722	1.724	0.213	0.211	0.946	1.000	1.716	0.210	0.209	0.944	1.000	1.025
High	**High**	**2.0**	1.586	1.588	0.231	0.229	0.949	1.000	1.575	0.229	0.228	0.949	1.000	1.015
High	**High**	**1.8**	1.584	1.589	0.233	0.229	0.940	1.000	1.577	0.231	0.228	0.943	1.000	1.018
High	**High**	**1.6**	1.587	1.586	0.230	0.229	0.948	1.000	1.575	0.228	0.227	0.947	1.000	1.017
High	**High**	**1.4**	1.585	1.583	0.231	0.229	0.949	1.000	1.574	0.229	0.226	0.948	1.000	1.019
High	**High**	**1.2**	1.589	1.585	0.231	0.229	0.947	1.000	1.577	0.228	0.226	0.944	1.000	1.025

Table 24: Simulation results under a large treatment effect and small sample sizes

[14] The full term for each statistical abbreviation in the headings of Table 22 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

[15]*Large treatment effect and medium sample sizes* (*250 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	1.719	1.719	0.066	0.066	0.949	1.000	1.717	0.069	0.068	0.950	1.000	0.919
Low	**Low**	**1.8**	1.719	1.717	0.067	0.066	0.945	1.000	1.717	0.069	0.068	0.946	1.000	0.955
Low	**Low**	**1.6**	1.719	1.717	0.066	0.066	0.950	1.000	1.717	0.067	0.067	0.951	1.000	0.974
Low	**Low**	**1.4**	1.717	1.717	0.066	0.066	0.949	1.000	1.717	0.067	0.067	0.950	1.000	0.987
Low	**Low**	**1.2**	1.717	1.716	0.066	0.066	0.950	1.000	1.716	0.066	0.066	0.950	1.000	0.996
Low	**High**	**2.0**	1.580	1.580	0.077	0.077	0.949	1.000	1.580	0.079	0.079	0.950	1.000	0.959
Low	**High**	**1.8**	1.580	1.577	0.077	0.077	0.950	1.000	1.577	0.079	0.078	0.947	1.000	0.963
Low	**High**	**1.6**	1.580	1.581	0.076	0.077	0.949	1.000	1.581	0.077	0.078	0.950	1.000	0.982
Low	**High**	**1.4**	1.581	1.581	0.077	0.077	0.955	1.000	1.581	0.077	0.077	0.952	1.000	0.987
Low	**High**	**1.2**	1.581	1.581	0.077	0.077	0.947	1.000	1.580	0.077	0.077	0.947	1.000	0.999
High	**Low**	**2.0**	1.723	1.724	0.096	0.096	0.951	1.000	1.712	0.096	0.095	0.947	1.000	1.001
High	**Low**	**1.8**	1.727	1.724	0.095	0.096	0.952	1.000	1.712	0.095	0.095	0.950	1.000	1.007
High	**Low**	**1.6**	1.725	1.722	0.096	0.096	0.946	1.000	1.712	0.096	0.095	0.945	1.000	1.016
High	**Low**	**1.4**	1.722	1.723	0.096	0.096	0.950	1.000	1.714	0.095	0.095	0.947	1.000	1.019
High	**Low**	**1.2**	1.723	1.726	0.097	0.096	0.949	1.000	1.718	0.096	0.095	0.947	1.000	1.027
High	**High**	**2.0**	1.585	1.587	0.103	0.103	0.953	1.000	1.574	0.102	0.103	0.951	1.000	1.008
High	**High**	**1.8**	1.586	1.584	0.103	0.103	0.948	1.000	1.573	0.103	0.103	0.947	1.000	1.012
High	**High**	**1.6**	1.586	1.585	0.104	0.104	0.946	1.000	1.574	0.103	0.103	0.945	1.000	1.018
High	**High**	**1.4**	1.586	1.586	0.104	0.103	0.946	1.000	1.577	0.103	0.102	0.946	1.000	1.024
High	**High**	**1.2**	1.587	1.588	0.106	0.104	0.943	1.000	1.580	0.105	0.102	0.944	1.000	1.023

Table 25: Simulation results under a large treatment effect and medium sample sizes

[15] The full term for each statistical abbreviation in the headings of Table 22 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

[16]*Large treatment effect and large sample sizes* (*500 subjects per treatment group*)

CS	TS	SDF	TTP	Latent					Observed					RE
				MTE	ESD	ASE	CP	P	MTE	ESD	ASE	CP	P	
Low	**Low**	**2.0**	1.718	1.717	0.046	0.047	0.958	1.000	1.716	0.047	0.048	0.957	1.000	0.930
Low	**Low**	**1.8**	1.718	1.718	0.047	0.047	0.948	1.000	1.718	0.048	0.048	0.945	1.000	0.945
Low	**Low**	**1.6**	1.715	1.717	0.047	0.047	0.945	1.000	1.717	0.048	0.048	0.950	1.000	0.970
Low	**Low**	**1.4**	1.719	1.718	0.048	0.047	0.946	1.000	1.718	0.048	0.047	0.944	1.000	0.986
Low	**Low**	**1.2**	1.717	1.718	0.047	0.047	0.950	1.000	1.718	0.047	0.047	0.951	1.000	1.000
Low	**High**	**2.0**	1.580	1.580	0.054	0.054	0.945	1.000	1.580	0.056	0.056	0.947	1.000	0.955
Low	**High**	**1.8**	1.581	1.581	0.055	0.054	0.948	1.000	1.581	0.056	0.055	0.949	1.000	0.969
Low	**High**	**1.6**	1.579	1.580	0.055	0.054	0.946	1.000	1.580	0.055	0.055	0.948	1.000	0.977
Low	**High**	**1.4**	1.582	1.578	0.055	0.054	0.947	1.000	1.578	0.055	0.055	0.948	1.000	0.986
Low	**High**	**1.2**	1.577	1.579	0.054	0.054	0.951	1.000	1.579	0.054	0.055	0.952	1.000	1.000
High	**Low**	**2.0**	1.725	1.723	0.068	0.068	0.949	1.000	1.710	0.068	0.068	0.943	1.000	1.004
High	**Low**	**1.8**	1.722	1.723	0.068	0.068	0.946	1.000	1.712	0.068	0.067	0.944	1.000	1.009
High	**Low**	**1.6**	1.722	1.723	0.068	0.068	0.951	1.000	1.712	0.067	0.067	0.949	1.000	1.016
High	**Low**	**1.4**	1.726	1.723	0.068	0.068	0.954	1.000	1.713	0.067	0.067	0.946	1.000	1.018
High	**Low**	**1.2**	1.722	1.724	0.069	0.068	0.947	1.000	1.716	0.068	0.067	0.946	1.000	1.023
High	**High**	**2.0**	1.585	1.586	0.073	0.073	0.948	1.000	1.573	0.073	0.073	0.947	1.000	0.999
High	**High**	**1.8**	1.587	1.587	0.074	0.073	0.949	1.000	1.575	0.073	0.073	0.947	1.000	1.018
High	**High**	**1.6**	1.584	1.587	0.075	0.073	0.951	1.000	1.577	0.074	0.073	0.945	1.000	1.012
High	**High**	**1.4**	1.585	1.585	0.073	0.073	0.951	1.000	1.575	0.072	0.073	0.947	1.000	1.021
High	**High**	**1.2**	1.585	1.586	0.073	0.073	0.951	1.000	1.578	0.072	0.073	0.950	1.000	1.018

Table 26: Simulation results under a large treatment effect and large sample sizes

[16] The full term for each statistical abbreviation in the headings of Table 22 are as follows:
CS – Control group skewness
TS – Treatment group skewness
SDF – Serial dilution factor
TTP – True treatment parameter
MTE – Mean log-transformed treatment estimate
ESD – Empirical standard deviation of treatment estimate
ASE – Average standard error of treatment estimate
CP – Coverage probability of treatment estimate
P – Power of simulation study
RE – Relative efficiency of treatment estimate comparing **latent titers** to **observed titers**

At small sample sizes, in comparing the impact of a large treatment effect to that of a trivial treatment effect (Table 24 to Table 7), with the exception of an increase to both the MTE and overall P, highly similar values and exact trends can be observed for the RE, MTE, ESD, ASE, and CP. In contrast to the setting with restricted ULQ, the RE increases at a much slower rate with smaller SDFs.

At medium sample sizes, in comparing the impact of a large treatment effect to that of a trivial treatment effect (Table 25 to Table 8), with the exception of an increase in both the MTE and overall P, the exact trends and highly similar values can be observed for the RE, MTE, ESD, ASE, and CP. The RE also shows a slower increasing trend as the SDF decreases.

At large sample sizes, in comparing the impact of a large treatment effect to that of a trivial treatment effect (Table 26 to Table 9), with the exception of an increase in both the MTE, and in P when the treatment group is highly skewed, the exact trends and highly similar values can be observed for the RE, MTE, ESD, ASE, and CP.

Appendix D: Extended SDF Impact on Proportional Odds Regression Statistics

Estimated standard deviation of the log-cumulative odds – ESD, $\sigma_e\left(\hat{\beta}_1\right)$

Figure 14: Scatter plots of ESDs from proportional odds regression to SDFs under different λ and β_1

The shapes and trajectories of the ESD plots for proportional odds regression as shown in Figure 14 mirror exactly to those of the ESD plots for simple linear regression in Figure 10.

Average standard errors of the log-cumulative odds – ASE, $\overline{\sigma}\left(\hat{\beta}_1\right)$

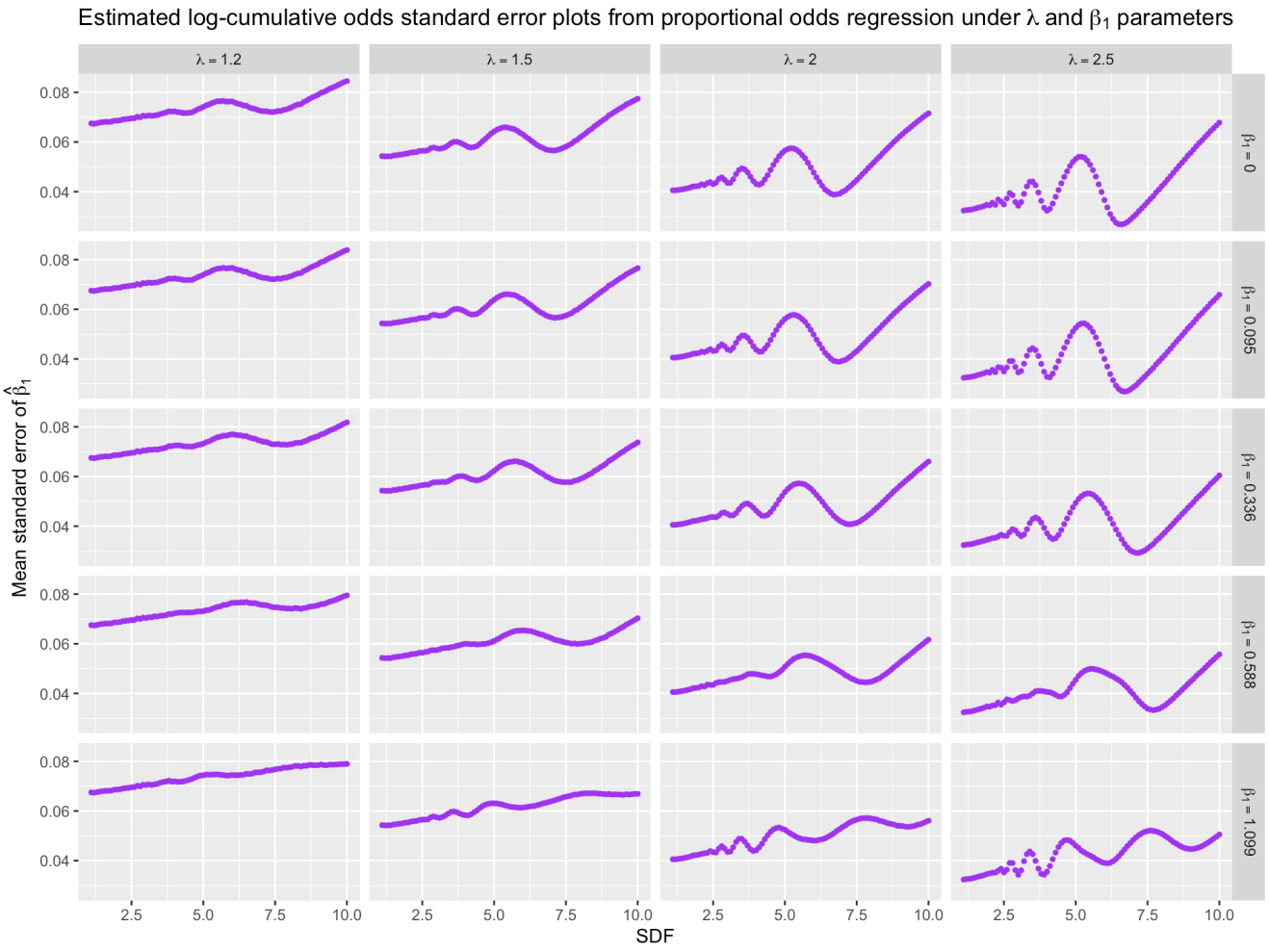

Figure 15: Scatter plots of ASEs from proportional odds regression to SDFs under different λ and β_1

The shapes and trajectories of the ASE plots for proportional odds regression as shown in Figure 15 mirror exactly to those of the ASE plots for simple linear regression in Figure 11.

Appendix E: R Program Summary

This appendix shows a summary of the R program used to conduct all the simulations in this study, as well as the R packages utilized.

```
## R version 4.0.5 (2021-03-31)
## Platform: x86_64-apple-darwin17.0 (64-bit)
## Running under: macOS Big Sur 10.16
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/
libRblas.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.0/Resources/lib/
libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## loaded via a namespace (and not attached):
##  [1] compiler_4.0.5    magrittr_2.0.1    tools_4.0.5       htmltools_0.5.1.1
##  [5] yaml_2.2.1        stringi_1.5.3     rmarkdown_2.7     knitr_1.33
##  [9] stringr_1.4.0     xfun_0.22         digest_0.6.27     rlang_0.4.10
## [13] evaluate_0.14
```