

Survival, Count and Discrete Mixture Models for Analyzing Tumor Latency and Multiplicity in Carcinogenesis and Chemoprevention Experiment

by

Rong Zhang

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
May 4, 2013

Keywords: Generalized Linear Model, Survival Analysis, Mixture model, zero-inflated models

Copyright 2013 by Rong Zhang

Approved by

Mark Carpenter, Chair, Professor of Mathematics and Statistics
Asheber Abebe, Associate Professor of Mathematics and Statistics
Peng Zeng, Associate Professor of Mathematics and Statistics

Abstract

Since the chemicals, including Bisphenol A (BPA), Bis(2-ethylhexyl) phthalate (DEHP), have been reported for potential carcinogens separately and genistein is an controversial chemopreventive isoflavone, animal experiment was designed and conducted for solving the following questions of interest: i) Can the compounds like BPA, DEHP and their combinations accelerate the carcinogenic process? ii) If yes, which one has significant stronger effect? iii) Does genistein have chemopreventive effect on those rats? And iv) What are the patterns when genistein was fed as a diet interacted with those treated compounds together?

Tumor latency, burden and multiplicity, analyzed univariately, were three issues of interests. The first two were survival time, and were analyzed by survivor functions estimated by nonparametric methods as well as parametrical survival regression analysis. Nonparametric methods did not detect any difference across ten groups while Weibull model fitted both of them better than Proportional hazards model. Tumor multiplicity was count data and analyzed by Poisson regression, negative binomial regression and their corresponding Zero-inflated models due to the problems of over-dispersion and excess zeros. The zero-inflated Poisson model reflected the multiplicity best yet negative binomial distribution fitted the data best among the four models.

The survival analysis and generalized linear modeling indicated that BPA, DEHP and their combinations decreased the time to first tumor but did not have significant effects on tumor burden and multiplicity. Genistein did not show any chemopreventive effect in this study no

matter solo or combined with other carcinogens. In application, this study indicates that early exposure to plastic consumer products which contain BPA and DEHP may increase the risk of breast cancer, while drinking soymilk may not help.

Acknowledgments

First of all, I would like to express my gratitude to my major professor Dr. Mark Carpenter for guiding me into this exciting research area of statistical inference methods in chemoprevention and carcinogenicity for my master program. Without his detailed guidance and helpful suggestions this thesis would not have been possible. I thank him for his patience and encouragement that carried me on through difficult times. I appreciate his vast knowledge and skills in many areas that greatly contributed to my research and this dissertation.

Special thanks are also given to the other members of my committee, Dr. Asheber Abebe and Dr. Peng Zeng, for their valuable assistance and comments. I am also benefited a lot from the courses taught by them.

I would like to extend my gratitude to all the professors in Dept. of Mathematics and Statistics who used to instruct my courses or helped me indirectly. My appreciation is also sent to Auburn University for the providing advanced learning environment for all the international students.

Finally, I am greatly indebted to my husband, parents and my parents-in-law, for their support, love, patience, and understanding.

Table of Contents

| | |
|---|------|
| Abstract | ii |
| Acknowledgments..... | iv |
| List of Tables | viii |
| List of Figures | ix |
| List of Abbreviations | x |
| Chapter 1: Introduction | 1 |
| 1.1 Generalized linear model | 3 |
| 1.1.1 General overview | 3 |
| 1.1.2 Poisson regression..... | 4 |
| 1.1.3 Zero-inflated count model..... | 7 |
| 1.1.4 Exponential regression..... | 9 |
| 1.2 Survival analysis | 10 |
| 1.2.1 Special features of survival data | 10 |
| 1.2.2 The survivor function..... | 11 |
| 1.2.3 The hazard function | 11 |
| 1.2.4 The median survival time..... | 12 |
| 1.2.5 Log-rank test | 13 |
| 1.2.6 Proportional hazards model | 13 |
| 1.2.7 Weibull model..... | 14 |

| | |
|--|----|
| 1.3 Literature review | 15 |
| 1.3.1 Tumor latency | 15 |
| 1.3.2 Tumor burden..... | 17 |
| 1.3.3 Tumor multiplicity..... | 17 |
| 1.4 Chemical compounds..... | 18 |
| 1.5 Motivation and goal | 20 |
| Chapter 2: Experiment design and methods | 22 |
| 2.1 Animals..... | 22 |
| 2.2 | 22 |
| 2.3 DMBA induced model..... | 23 |
| 2.4 Statistical methods | 23 |
| Chapter 3: Results | 24 |
| 3.1 Tumor latency | 24 |
| 3.1.1 General description | 24 |
| 3.1.2 Survival analysis | 26 |
| 3.1.3 Model fitting | 28 |
| 3.2 Tumor burden..... | 32 |
| 3.2.1 General description | 32 |
| 3.2.2 Survival analysis | 34 |
| 3.2.3 Model fitting | 36 |
| 3.3 Tumor multiplicity | 40 |
| 3.3.1 General description | 40 |
| 3.3.2 Poisson regression..... | 42 |

| | |
|--|----|
| 3.3.3 Negative binomial regression | 43 |
| 3.3.4 Zero-inflated Poisson regression..... | 44 |
| 3.3.5 Zero-inflated negative binomial regression | 46 |
| 3.3.6 Model comparison | 48 |
| Chapter 4: Discussion | 50 |
| 4.1 Interpretation..... | 50 |
| 4.2 Compare to Previous literature | 56 |
| 4.3 limitations of the study..... | 57 |
| Chapter 5: Conclusion..... | 58 |
| 5.1 Implications of results..... | 58 |
| 5.2 Future research..... | 58 |
| Reference | 60 |
| Appendix 1: Tables and Figures | 65 |
| Appendix 2: Partial SAS Code | 69 |

List of Tables

| | |
|--|----|
| Table 1: Descriptive Table for Tumor Latency | 25 |
| Table 2: Estimates for Tumor Latency Sored by Median from LifeTest procedure | 28 |
| Table 3: Estimates of Final Proportional Hazards Model for Tumor Latency | 30 |
| Table 4: Estimates of Final Weibull Model for Tumor Latency | 31 |
| Table 5: Model Selection for Weibull Distribution for Tumor Latency..... | 32 |
| Table 6: Descriptive Table for Tumor Burden | 33 |
| Table 7: Estimates for Tumor Burden by Each Group from LifeTest | 35 |
| Table 8: Estimates of Weibull Model for Tumor Burden | 39 |
| Table 9: Model Selection for Weibull Distribution for Tumor Burden..... | 40 |
| Table 10: Descriptive Table for Tumor Multiplicity Sored by Diet and Treatment..... | 41 |
| Table 11: Final Poisson Model for Tumor Multiplicity..... | 43 |
| Table 12: Negative Binomial Model for Tumor Multiplicity..... | 44 |
| Table 13: ZIP Model for Tumor Multiplicity | 45 |
| Table 14: ZINB model for tumor multiplicity | 47 |
| Table 15: Model Comparison for Four Full Models..... | 49 |
| Table 16: Model Selection of Poisson Model for Tumor Multiplicity | 65 |
| Table 17: Model Selection of NB Model for Tumor Multiplicity | 65 |
| Table 18: Variable Selection for Poisson Given in ZIP Model for Tumor Multiplicity | 65 |
| Table 19: Variable Selection for NB Given in ZINB Model for Tumor Multiplicity | 66 |

List of Figures

| | |
|--|----|
| Figure 1: Boxplot for Tumor Latency by Ten Groups | 26 |
| Figure 2: Kaplan-Meier Curves for Tumor Latency by Ten Groups | 27 |
| Figure 3: Negative log of Survivor function Plots for Tumor Latency by Ten Groups | 29 |
| Figure 4: Log of Negative Log of Survivor Functions Plots for Tumor Latency by Ten Groups..... | 29 |
| Figure 5: Boxplot for Tumor Burden by Ten Groups | 34 |
| Figure 6: Kaplan-Meier Curves for Tumor Burden by Ten Groups | 36 |
| Figure 7: Negative log of Survivor function for Tumor Burden by Ten Groups | 37 |
| Figure 8: Log of Negative Log of Survivor Functions for Tumor Burden by Ten Groups..... | 38 |
| Figure 9: Boxplot for Tumor Multiplicity by Ten Groups | 42 |
| Figure 10: Model Comparisons for Fitted and Observed Models of Tumor Multiplicity | 49 |
| Figure 11: Boxplot for Tumor Latency by Treatment, Adjusting Diet..... | 66 |
| Figure 12: Boxplot for Tumor Burden by Treatment, Adjusting Diet..... | 67 |
| Figure 13: Boxplot for Tumor Multiplicity by Treatment, Adjusting Diet | 68 |
| Figure 14: Histogram of Overall Tumor Multiplicity..... | 69 |

List of Abbreviations

| | |
|--------|--|
| BPA | Bisphenol A |
| CDF | Cumulative Distribution Function |
| CON | Control |
| DMBA | Dimethylbenzanthracene |
| DEHP | Di-2-ethylhexyl phthalate |
| F | Failure |
| GEN | Genistein |
| GENMOD | Generalized linear models procedure in SAS |
| GLM | Generalized Linear Models |
| HC | High Combination |
| HZF | Hazard Function |
| LC | Low Combination |
| NB | Negative Binomial |
| P | Probability |
| PDF | Probability Density Function |
| PHM | Proportional Hazards Model |
| PMF | Probability Mass Function |
| S | Survival |
| SO | Sesame Oil |

| | |
|------|---------------------------------|
| SVF | Survivor Function |
| ZINB | Zero-inflated Negative Binomial |
| ZIP | Zero-inflated Poisson |

Chapter 1: Introduction

This is an applied statistical thesis that discusses the survival, count and discrete mixture models for tumor latency, tumor burden and tumor multiplicity for DMBA induced mammary tumor in carcinogenesis and chemoprevention research. Since the chemicals, including Bisphenol A (BPA), Bis(2-ethylhexyl) phthalate (DEHP), have been reported for potential carcinogens separately and genistein is an controversial chemopreventive isoflavone, animal experiment was conducted for solving the following questions of interest: i) Can the compounds like BPA, DEHP and their combinations, which were orally exposed to the female offspring of the DMBA induced rats accelerate the carcinogenic process? ii) If yes, which one has significant stronger effect? iii) Does genistein have chemopreventive effect on those rats? And iv) What are the patterns when genistein was fed as a diet interacted with those treated compounds together?

Correspondingly, if we convert these biological questions into statistical one, our interested problems would be: i) Could the chemical treatments and genistein diet explain the variability of the tumor latency, burden and multiplicity? ii) What kind of distributions did them follow univariately?

In order to answer these specific questions, a collection of statistical approaches were applied to model the data and explain the variability. The multivariate data analysis did not developed in this study. Instead, we analyzed Tumor latency, tumor burden and tumor multiplicity separately as univariate response as we want to consider each carcinogenic and chemopreventive result separately. These three variables stood out as the best choices for the responses not only because of their popularity among researchers who study carcinogenesis and chemoprevention, but also because they are the keys for evaluating the process of

carcinogenicity and the results of chemoprevention. Although there were various definitions about these three terms, in this study, we define tumor latency as the time between the date that the first tumor had been detected and the date of birth for each rat. Similarly, tumor burden, being defined as the time from birth to sacrifice given tumor onset, is the survival time as well in this study. Tumor multiplicity was defined as the number of the total tumor per rat.

Tumor latency and burden are typical survival time with censored data so that can be modeled appropriately by survival analysis. Tumor multiplicity, the number of tumors per animal at a particular time was expected to follow a Poisson distribution [1] so that can be modeled as Poisson regression. However, the negative binomial (NB) distribution, which is well known to be a gamma mixture of Poisson, is much more widely used in modeling over-dispersed count data [2]. In addition, the zero fraction of tumor multiplicity tended to be large according to histogram plot and descriptive table in chapter 3. Therefore, negative binomial regression were applied to address over-dispersion issue and zero-inflated count models, including zero-inflated Poisson model and zero-inflated negative binomial model were fitted to balance excess zeros.

In sum, the statistical methods applied in this study were all fallen into the big umbrella of generalized linear model, which is including survival analysis and count data regression, only except the zero-inflated count models. They are, however, very flexible extensions for general Poisson regression.

This introduction chapter focused on the methodology background of generalized linear model, survival analysis, and zero-inflated count model. It also summarized statistical inference methods on tumor latency, tumor burden and tumor multiplicity, and illustrated the motivation of

this study with more details. Throughout the paper, capital letter represents random variable, and boldface notation is used to denote matrices and vectors in the equations and other statements.

1.1 Generalized linear model

Poisson regression is a form of regression analysis used to model count data and contingency tables, and Exponential regression can be used to model survivor function in survival data analysis. Both of the distributions belong to the exponential families and can be modeled by generalized linear model. Therefore, a brief introduction for generalized linear model is present at the beginning.

1.1.1 General overview

The term ‘Generalized linear model (GLM)’ is due originally to Nelder and Wedderburn (1972)[3], who showed how linearity could be exploited to unify apparently diverse statistical technique. In this section, the notation and approach are closely following the book *Generalize Linear Model* written by McCullagh and Nelder (1989) [4]. Begin by considering the normal linear regression model,

$$Y_i = \mathbf{X}_i\boldsymbol{\beta} + \varepsilon_i, i = 1, \dots, n, \quad (1.1)$$

where Y_i is a dependent variable, \mathbf{x}_i is a vector of k independent variables, $\boldsymbol{\beta}$ is a k -by-1 vector of unknown parameters and the ε_i are normally distributed random errors with zero-mean and constant variance σ^2 . To extend to GLM, it allows for response variables that have other than a normal distribution and allows modeling some function of the mean.

All GLMs have three components [5]:

1) The *random component* identifies the response variable \mathbf{Y} and assumes a probability distribution for it. The distribution may be any probability distribution that comes from an exponential family other than normal.

2) The *systematic component* specifies the explanatory variables for the model. Covariates $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ produce a linear predictor $\boldsymbol{\eta}$ given by $\boldsymbol{\eta} = \sum_1^p \mathbf{x}_j \beta_j$. This is the only component that does not change when it extends to GLM.

3) The *link function* specifies a function of the expected value (mean) of \mathbf{Y} , which the GLM relates to the explanatory variables through a prediction equation having linear form. If we write $\eta = g(\mu)$, then the $g(\cdot)$ will be called the link function.

In this study, one of the responses, tumor multiplicity, would be assumed to follow the Poisson distribution instead of normal distribution. Meanwhile, other two responses, time to first tumor (tumor latency) and time to death (tumor burden) will be assumed to follow the exponential distribution in GLM. Therefore, these two kinds of regression will be discussed with more details in the next two sections.

1.1.2 Poisson regression

Poisson regression models are generalized linear models with the logarithm as the link function, and the Poisson distribution function. The PMF of Poisson distribution is given by

$$P(X=x | \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x=0, 1, 2 \dots \quad (1.2)$$

The assumptions of Poisson regression include:

1. Logarithm of the response rate changes linearly with equal increment increases in the predictors.
2. Changes in the rate from combined effects of different predictors are multiplicative.
3. At each level of the covariates the number of cases has variance equal to the mean.
4. Observations are independent.

Therefore, the expected mean λ is a loglinear function of predictors.

$$\log \lambda_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}. \quad (1.3)$$

The unknown parameters can be estimated by the maximum likelihood estimation, which is more convenience to calculate the log-likelihood function. Once the third assumption is violated, which means the variance is larger than the mean, it is known as extra-Poisson variation, or over-dispersion. A common cause of this problem is heterogeneity among subjects. One simple way to correct over-dispersion is correct the standard errors and chi-square if the lack of efficiency of conventional estimates can be simply ignored [6]. Under some circumstances, the problem of over-dispersion can also be solved by fitting a negative binomial regression model instead [7].

The negative binomial is another distribution that is concentrated on the nonnegative integers. Unlike the Poisson, it has an additional parameter such that the variance can exceed the mean. Indeed the negative binomial distribution is widely used to model count data [4] and the PMF is

$$P(X=x | r, \lambda) = \frac{\lambda^x}{x!} \frac{\Gamma(r+x)}{\Gamma(r)\Gamma(r+x)} \frac{1}{(1+\frac{\lambda}{r})^r}, x = 0, 1, 2, \dots \quad (1.4)$$

with mean λ and variance $\lambda (1 + \frac{\lambda}{r})$ for $r \geq 0$, where r is the number of failures until the experiment is stopped and the shape parameter of the gamma mixing distribution as well. When the over-dispersion parameter $\alpha = \frac{1}{r}$ goes to zero, the second factor will converge to one, and the third to the exponent function $e^{-\lambda}$. So the negative binominal distribution corresponds to the Poisson distribution [8]. So it is a generalization of the Poisson model, and its loglinear function can be written as

$$\log \lambda_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \frac{1}{r} \varepsilon_i, \quad (1.5)$$

where the dependent variable Y_i is assumed to have a Poisson distribution with the expected value λ_i , conditional on ε_i . The $\exp(\varepsilon_i)$ is assumed to have a standard gamma distribution, following that the unconditional of Y_i is a negative binomial distribution. Equation 1.5 can be efficiently estimated by maximum likelihood estimation method as well.

The Poisson and negative binomial distributions are well suited for describing counts of events that occur in some interval of time. If the events are counted over different lengths of time for different individuals, there is a need for standardization. The time incorporated model is given by

$$P(X=x | \lambda, t) = \frac{e^{-\lambda t} \lambda t^x}{x!}, \quad x=0, 1, 2 \dots \quad (1.6)$$

where t is the length of the observation interval for individual, and λt is the expected value. The loglinear function in this case is

$$\log E(Y_i) = \log (\lambda_i t_i) = \log t_i + \log \lambda_i = \log t_i + \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} \quad (1.7)$$

It implies that the logarithm of the observation time should be on the right-hand side of the equation, with a coefficient of 1.0. Therefore, the logarithm of tumor burden had been set as an offset variable in this study.

1.1.3 Zero-inflated count model

Another common problem with Poisson regression is excess zeros. For example, if there are two processes at work, one determining whether there are zero events or any events, and a Poisson process determining how many events there are, there will be more zeros than a Poisson regression would predict. In this case, the zero-inflated count can be applied to take care of the excess zero-count data in unit time [9].

The main motivation for zero-inflated count models is that real-life data frequently display over-dispersion and excess zeros. These distributions are a mixture of two distributions, a degenerate component that is zero with certainty and a second component that includes zeros and positive values (e.g., the Poisson distribution). The general form of these distributions is

$$P(Y=y) = \begin{cases} \pi + (1 - \pi)f(0), & y = 0, \\ (1 - \pi)f(y), & y = 1, 2, \dots \end{cases} \quad (1.8)$$

where π is the probability that an observation comes from the degenerate component. Two commonly used zero-inflated distributions are the zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) [10].

In order to describe ZIP model explicitly, assume that the second component which includes zeros and positive integer is Poisson distribution with mean and variance λ . Therefore, the PMF of ZIP model is

$$P(Y=y) = \begin{cases} \pi + (1 - \pi)e^{-\lambda}, & y = 0, \\ (1 - \pi)e^{-\lambda} \frac{\lambda^y}{y!}, & y = 1, 2, \dots \end{cases} \quad (1.9)$$

where π is the probability that an observation comes from the degenerate component. The mean and variance of the ZIP model is given by

$$E(Y) = E(E(Y|X)) = E(X \lambda) = (1 - \pi) \lambda,$$

$$\text{and var}(Y) = \text{var}(E(Y|X)) + E(\text{var}(Y|X))$$

$$= \text{var}(X \lambda) + E(X \lambda)$$

$$= \lambda^2 \pi (1 - \pi) + (1 - \pi) \lambda$$

$$= \lambda (1 - \pi) (1 + \lambda \pi)$$

ZINB model is a more realistic model for application. Assume that the second component which includes zeros and positive integer is a gamma distribution with mean μ and heterogeneity parameter α . Therefore, the PMF of ZINB model is

$$P(Y=y) = \begin{cases} \pi + (1 - \pi) \left(\frac{\alpha}{\alpha + \mu}\right)^\alpha, & y = 0, \\ (1 - \pi) \frac{\Gamma(y + \alpha)}{\Gamma(\alpha) \Gamma(y + 1)} \left(\frac{\alpha}{\alpha + \mu}\right)^\alpha \left(\frac{\mu}{\alpha + \mu}\right)^y, & y = 1, 2, \dots \end{cases} \quad (1.10)$$

where π is the probability that an observation comes from the degenerate component. Similarly, the mean and variance of the ZINB model is given by

$$E(Y) = E(E(Y|X)) = E(X \lambda) = (1 - \pi) \mu,$$

$$\text{and var}(Y) = \text{var}(E(Y|X)) + E(\text{var}(Y|X))$$

$$= \text{var}(X \mu) + E\left(X \left(\mu + \frac{1}{\alpha} \mu^2\right)\right)$$

$$\begin{aligned}
&= \mu^2 \pi (1-\pi) + (1-\pi) \left(\mu + \frac{1}{\alpha} \mu^2 \right) \\
&= \mu (1-\pi) \left(1 + \frac{1}{\alpha} \mu + \pi \mu \right)
\end{aligned}$$

Note that these two zero-inflated models have the similar mean variance property as the standard negative binomial models. It is not surprising that in some cases these models are confounded with each other and all of them can fit the data well. General, the ZINB model seems to be most flexibility and efficiency but has very limited application due to the large estimated variance and large AIC/BIC [11].

1.1.4 Exponential regression

Exponential regression models are generalized linear models with the inverse as the link function, and the Exponential distribution function. The PDF for exponential distribution is given by

$$P(X=x | \lambda) = \lambda e^{-\lambda x}, \quad x \geq 0 \tag{1.11}$$

with mean $1/\lambda$ and variance $1/\lambda^2$. The exponential distribution occurs naturally when describing the lengths of the inter-arrival times in a homogeneous Poisson process. Even though it can be used to model survival data, the goodness of fit test may be failed due to the large deviation.

Weibull distribution is another distribution related to both the exponential and gamma families. If $X \sim \text{exponential}(\beta)$, then $Y = X^{1/\gamma}$ has a Weibull (γ, β) distribution [1]. It plays a central role in survival analysis as the normal distribution dose in the standard analysis [12]. More details about Weibull model will be provided at subsection 1.2.7., under the topic of survival analysis.

1.2 Survival analysis

Survival analysis is the phrase used to describe the analysis of data that correspond to the time from a well-defined time origin until the occurrence of some particular event or endpoint [12]. This method has been widely used in medical research and is also called reliability analysis in engineering. The typical questions that can be answered by survival analysis are: i) What is the probability of a population which will survive past a certain time? ii) Which population is more likely to have a longer lifespan and better prognosis, when there are several treatments applied? iii) Can multiple causes of death or failure be taken into account? And iv) Which kind of cause should be taken into account? In this section, the notation and approach are closely following the book *Modelling survival data in medical research* written by D. Collett (1994) [12].

1.2.1 Special features of survival data

There are two special features that determine the differences between analyzing survival data and the standard data. The first one is that the survival data are generally not symmetrically distributed [12]. The histogram of survival time tends to be positively skewed, which means it has a longer tail to the right of the median.

The second feature is that the survival data are frequently censored. Censoring occurs when the endpoint of interest has not been observed for that individual. If the actual unknown survival time is larger than observed survival time, which is also the most commonly case, it is called right censoring. For instant, one is still alive at the endpoint of the research. The information contained by this individual is not complete but still important. On the other hand, if the actual unknown survival time is shorter than observed survival time, it is known as left censoring. For example, the cancer has already recurred when the patient is examined. The actual

time for recurrence was not observed and the only thing we know is it shorter than the examination time. In addition, when individuals are known to have experienced a failure within an interval of time, it is called interval censoring.

1.2.2 The survivor function

The survivor function (SVF), $S(t)$, captures the probability that the survival time is greater than or equal to a specified time t . The mathematical equation for SVF is given by

$$S(t) = P(T \geq t) = 1 - F(t), \quad (1.12)$$

and represents the probability that an individual survives from the time origin to some time beyond t . Every SVF is monotonically decreasing function, with right continuous property.

There are several methods to estimate survivor function, such as Life-table estimate (for non-censored data), Kaplan-Meier estimate (for censored data), and so on.

Kaplan-Meier estimate of the survivor function is also known as the product-limit estimate of the survivor function. By defining n_j as the number of successes including the number of about to fail before time t_j and d_j as the number of failure at time t_j , the K-M estimate is given by,

$$S(t) = \prod_{j=1}^k \left(\frac{n_j - d_j}{n_j} \right), \quad k=1, 2, \dots, r \quad (1.13)$$

for $t_{(k)} \leq t < t_{(k+1)}$, with $S(t) = 1$ for $t < t_{(1)}$ and where $t_{(r+1)}$ is taken to positive infinite. In this study, Kaplan-Meier estimate was applied in the chapter 3.

1.2.3 The hazard function

The hazard function (HZF), $h(t)$, is another important function of central interest. It is the probability that an individual dies at time t , conditional on he or she having survived to that time

[12]. HZF represents the instantaneous death rate for an individual surviving to time t . The mathematical equation for HZF is given by

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \right\}. \quad (1.14)$$

The relationship between SVF and HZF can be derived from equation (1.12). The limiting value of instantaneous failure time is the derivative of $F(t)$ with respect to t , denoted by $f(t)$. So

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} [\ln S(t)]. \quad (1.15)$$

1.2.4 The median survival time

The median survival time is the preferred summary measure of the location since the distribution tends to be positively skewed. It is straightforward to obtain once the SVF has been estimated. The median survival time is the 50th percentile of the distribution, and is given by the value $t(50)$ which is such that $S\{t(50)\} = 0.5$. Since the non-parametric estimates of $S(t)$ are step-function, it will not usually be possible to estimate the survival time that makes the SVF exactly equal to 0.5. Therefore, the smallest observed survival time for which the value of the estimated SVF is less than or equal to 0.5 is defined as the estimated median survival time. In mathematical terms,

$$\hat{t}(50) = \min \{t_i \mid \hat{S}(t_i) \leq 0.5\}, \quad (1.16)$$

where t_i is the observed survival time for the i 'th individual, $i = 1, 2, \dots, n$ [12].

1.2.5 Log-rank test

The log-rank test is a hypothesis test to compare the survival distributions of two samples, which also can be extended to do multiple comparisons. It is a nonparametric test and fit the survival data by natural. The test was first proposed by Nathan Mantel and was named the log-rank test by Richard and Julian Peto.[10, 11], and can also be viewed as a time stratified Mantel-Haenszel test. The log-rank test is a large-sample chi-square test that uses as its test criterion a statistic that provides an overall comparison of the Kaplan-Meier curves being compared. This (log-rank) statistic, like many other statistics used in other kinds of chi-square tests, makes use of observed versus expected cell counts over categories of outcomes. The categories for the log-rank statistic are defined by each of the ordered failure times for the entire set of data being analyzed.

Let O_{1j} and O_{2j} represent observed number of events in group 1 and group 2, respectively. And $e_{1j} = n_{1j}O_j / n_j$, $e_{2j} = n_{2j}O_j / n_j$ be the expected number of events in each group, where n_{ij} is the total number of each group and n_j is the total number for both group. So that we have

$$\chi^2 = \frac{[\sum_{j=1}^r (O_{1j} - e_{1j})]^2}{\text{Var}(\sum_{j=1}^r (O_{1j} - e_{1j}))} \sim \chi_1^2. \quad (1.17)$$

The null hypothesis is that there is no difference in the survival experiences in the two groups. If the value of χ^2 is larger than the critical value, the null hypothesis will be rejected.

1.2.6 Proportional hazards model

An approach based on statistical modeling can be used in order to explore the relationship between survival time and explanatory variables. The most commonly used model is the proportional hazards model, which was proposed by Cox in 1972 and has also be known as the

Cox regression model [13]. The model is only based on proportional hazards assumption, which is that hazard functions are proportional over time. There is no particular distribution is assumed, especially for base line distribution. Therefore, it is a semi-parametric model.

Proportional hazards model can be expressed in the form

$$h_N(t) = \psi h_S(t), \quad (1.18)$$

where ψ is a constant, known as the relative hazard or hazard ratio, and $h_S(t)$ and $h_N(t)$ are the hazards of death at time t for patients on the standard treatment and new treatment, respectively [12]. Since the hazard ratio cannot be negative, the general proportional hazards function can be written as

$$h_i(t) = \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}) h_0(t) \quad (1.19)$$

where β_i s are the coefficients of linear component. This is a generalized situation where the hazard of death at a particular time depends on the explanatory variables $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$. There are two types of variable on which a hazard function may depend, namely variates and factors. A variate is often a continuous variable such as height, weight, and age, which takes numerical values. A factor is a categorical variable like gender which takes a limited set of values.

The maximum likelihood estimation can be used to fit the proportional hazards model in order to estimate the unknown coefficients and hazard ratios. The maximization is generally accomplished using the Newton-Raphson procedure.

1.2.7 Weibull model

The inferences based on specific distribution assumption will be more precise if the assumption is valid. In practice, the assumption of a constant hazard function or equivalently of

exponentially distributed survival times is rarely tenable. A more general form of hazard function is such that

$$h(t) = \lambda r t^{r-1}, \quad (1.20)$$

for $0 \leq t < \infty$, $\lambda > 0$, $r > 0$ where λ is the scale parameter and r is the shape parameters on which the model depended. In particular case, when $r=1$, it becomes exponential distribution. For other r values, the HZF is a monotone function. The generalized SVF is given by

$$S(t) = \exp(-\lambda t^r). \quad (1.21)$$

The corresponding PDF is

$$f(t) = \lambda r t^{r-1} \exp(-\lambda t^r), \quad (1.22)$$

for $0 \leq t < \infty$, with scale parameter λ and shape parameter r .

1.3 Literature review

In this section, the review was focused on the concepts and statistical inference methods on tumor latency, tumor burden and tumor multiplicity. These three topics were chosen into this thesis, not only because of their popularity among researchers who study chemoprevention, but also because they are the key points to evaluate the development of tumor and the results of chemoprevention. A brief review for the chemical compounds that be used in the study was included in this section as well.

1.3.1 Tumor latency

Tumor latency, or dormancy, is a well-recognized clinical phenomenon and induction or maintenance of this state would appear to offer a novel therapeutic approach to limiting the effects of neoplastic disease [14]. In the clinic, tumor dormancy is observed in local recurrences

or metastases. It usually refers to the time after treatment that a patient is asymptomatic but still carries local remnant or disseminated tumor cells that do not grow into overt lesions [15]. In pathological, it means the time that allows the presumably malignant cells, shed from the primary mass, increase in population size to become a detectable recurrence [14]. The mechanism is imperfectly understood and the definition is still vague. In this study, we define tumor latency as the time between the date that the first tumor had been detected and the date of birth for each rat.

There are evidences indicating that the cancer recurrence after therapy and long periods of remission is frequent. For example, 20-45% of patients with breast or prostate cancer will relapse years or decades later [16-18]. Tumor latency can be considered as a component of cancer progression and can be suppressed if the chemoprevention is applied right in time [19]. Therefore, how to prolong the tumor latency becomes a hot topic in oncology and pathology.

Statistically, there are nonparametric and parametric ways available to analyze tumor latency. The Kaplan-Meier curve is a widely used nonparametric method to estimate the survival curve when the first tumor occurrence is defined as the endpoint event [20-22]. In addition, the following methods are popular in addressing tumor latency. Proportional hazards regression is a semi-parametric way, while the Weibull, log-logistic and other distributions for survival data are parametric methods [12, 23]. The software tools available for implementing the survival analysis method include the LIFETEST and LIFEREG procedure in Statistical Analysis Software (SAS). They are available to estimate the survivor function, to run Log-rank test, and to setup the survival model for tumor latency [20, 22, 24].

1.3.2 Tumor burden

There are many definitions for tumor burden with respect to different perspectives and disciplines [25-27]. In general, tumor burden also called tumor load, referring to the number of cancer cells, the size of a tumor, or the amount of cancer in the body. In this study, the mice were designed to be sacrificed when the burden is exceeded 10% of body weight. Hence the time from birth to sacrifice given tumor onset is defined as tumor burden in this thesis. In other words, tumor burden is the survival time in this study. It is worth to note that setting the endpoint as sacrificing instead of natural death might introduce some bias to this study.

The statistical inference methods for analyzing tumor burden or survival time are practically the same set with those for tumor latency because of the same or similar data properties. The method of Kaplan -Meier [28] can be used to obtain survival and relapse-free survival curves [25]. Survival times and disease-free survival times in distinct patient groups can be compared using the generalized Wilcoxon test [29]. The proportional hazards model of Cox can be used to detect explanatory variables' impact on tumor latency and tumor burden [30].

1.3.3 Tumor multiplicity

Tumor multiplicity (either total or malignant tumors only) , defined by Waalkes 2004, is the number of tumors per rat either of a particular organ or inclusively at any site [31]. In this study we define tumor multiplicity as the number of the total tumor per rat. It has been shown that the development of a tumor is a rare event occurring among a large population of cells at risk [32]. Therefore, the number of tumors per animal at a particular time after treatment might be expected to follow a Poisson distribution [33]. One implicit assumption of using Poisson distribution is that, the mean and variance of the number of tumors per animal for this distribution are supposed both equal to λ . However, usually even in highly controlled

experiments, the assumption of equal mean and variance does not hold true. Due to various sources of variation in the stochastic process of tumor formation, the data usually exhibit significantly bigger variance than the mean, which is known as over-dispersion [34-37]. Therefore, negative binomial model might be a better distribution to represent the data by using two parameters—mean and exponent determined by the interanimal homogeneity of tumor response [38]. In reality the negative binomial distribution has been widely used to model count data [4]. When the over-dispersion parameter k goes to 0, it corresponds to the Poisson distribution [8].

Dunson (1999) described that the zero-inflated Poisson model was a method for modeling carcinogenicity from animal studies where the data consist of counts of the number of tumors present over time [39]. The method is applied to testing for a dose-related trend in both tumor incidence and multiplicity in carcinogenicity studies [39]. Kim (2010) performed zero-inflated negative binomial distribution for modeling tumor multiplicity to reflect a high zero count [40].

Besides modeling, some of researchers were interested in the trend in proportion of tumor multiplicity. Cochran-Armitage test, named for William Cochran and Peter Armitage, was being widely used for this purpose [22, 24, 41]. In addition, others in carcinogenicity used Tukey-Kramer test or Dunnett's multiple comparisons in ANOVA to compare the difference of tumor multiplicity across groups [42-44].

1.4 Chemical compounds

This was an application study which was focus on statistical inference methods and models but also involved several chemicals that was introduced as predictors / independent

variables in the model. Therefore, a brief introduction and review for those chemicals will provide a better understanding background.

7, 12-dimethylbenz (a) anthracene (DMBA), an organ- and site specific procarcinogen, requires metabolic activation to become an ultimate carcinogen. The dihydrodiol epoxide (ultimate carcinogen) and other toxic reactive oxygen species formed during its metabolic activation can cause chromosomal damage by binding with adenine residues of DNA, contributing to mutagenesis and carcinogenesis [45]. DMBA can also interact with the estrogen receptor and partially mimic both the positive and negative feedback actions of estradiol in ovariectomized rats [46].

Bisphenol A (BPA) is a synthetically produced chemical used in the manufacture of polycarbonate plastics and epoxy resins. These products are utilized in a plethora of commonly used consumer goods, such as food and beverage containers, the lacquer lining of canned foods and drinks, infant formula bottles, receipts, office water cooler tanks, laboratory and hospital supplies, and some dental sealants [47]. BPA is reported to alter normal estrogen, androgen, and thyroid hormone signaling in vitro. It has been shown to cause adverse effects in breast cancer cell lines, including the induction of cell proliferation, producing oxidative stress, and altering cell signaling pathways involved in carcinogenesis and glucose homeostasis. Further research has shown that BPA is capable of antagonizing the cytotoxic effects of certain chemotherapeutics, such as doxorubicin, cisplatin, and vinblastine [48].

Bis(2-ethylhexyl) phthalate (DEHP) is an organic compound that is the most common of the class of phthalate plasticizers. DEHP has a low vapor pressure, but the temperatures for processing PVC articles are often high, leading to release of elevated levels, raising concerns

about health risks. It can be absorbed from food and water. Higher levels have been found in milk and cheese. Reports have been shown that DEHP is a potential endocrine disruptor. It has testicular toxicity [49], Ovarian toxicity [50], developmental toxicity [51], nephrotoxicity [52] and other toxicities.

Genistein is one of several known isoflavones. Epidemiology reports indicate that women consuming diet high in soy products, containing large amounts of phytoestrogens especially genistein, have a low incidence of breast cancer [53, 54]. Genistein (major isoflavone in soya) has been reported to have weak estrogenic and anti-estrogenic properties [55, 56], to be an antioxidant [57] and to inhibit the protein tyrosine kinase and topoisomerase II activity [58, 59] and also angiogenesis [60]. It is interesting that animals treated neonatally and prepubertally with genistein have reduced incidence and multiplicity of 7,12-dimethylbenzanthracene (DMBA)-induced mammary adenocarcinomas, decreased numbers of terminal end buds but increased numbers of lobular structures [61-63]. On the contrary, in adult animals, the effects of genistein on mammary tumor development remain controversial. For example, in adult mice, genistein increases mammary tumorigenesis [64, 65].

1.5 Motivation and goal

According to the literature view, Bisphenol A (BPA), exhibiting hormone-like properties, has raised concerns about its suitability in consumer products and food containers. Bis(2-ethylhexyl) phthalate (DEHP), most common plasticizer, is also reported to be a potential endocrine disruptor. Genistein, a potential chemopreventive isoflavone, is still controversial yet. Therefore, a highly controlled animal experiment was designed and conducted in order to find more evidence for the carcinogenesis or chemopreventive effect of these chemicals, and to explore more information, such that which one has stronger carcinogenic effect if they are all

carcinogens, and what the result would be if potential chemopreventive chemical genistein combined with these carcinogens. In this case, we have four specific target questions need to answer, which are i) Can the compounds like BPA, DEHP and their combinations, which were orally exposed to the female offspring of the DMBA induced rats accelerate the carcinogenic process? ii) If yes, which one has significant stronger effect than others? iii) Does genistein have chemopreventive effect on those rats? And iv) What are the patterns when genistein was fed as a diet interacted with those treated compounds together?

According to the natural of the dataset, two general statistical question can cover those four biological questions. Therefore, our interested problems would become: i) Could the chemical treatments and genistein diet explain the variability of the tumor latency, burden and multiplicity? ii) What kind of distributions did them follow univariately?

In order to answer these specific questions, a collection of statistical approaches were applied to model the data and explain the variability. The multivariate data analysis did not applied in this study. Instead, we analyzed Tumor latency, tumor burden and tumor multiplicity separately as univariate response as we want to consider each carcinogenic and chemopreventive result separately. Although there were various definitions about these three terms, in this study, we define tumor latency as the time between the date that the first tumor had been detected and the date of birth for each rat. Similarly, tumor burden, being defined as the time from birth to sacrifice given tumor onset, is the survival time as well in this study. Tumor multiplicity was defined as the number of the total tumor per rat.

Obviously, tumor latency and burden are typical survival time with censored data so that can be modeled appropriately by survival analysis. Nonparametric methods as well as

parametrical survival regression analysis will be applied to analyze them. Tumor multiplicity, the number of tumors per animal at a particular time was expected to follow a Poisson distribution so that can be modeled as Poisson regression. However, if the data has over-dispersion problem, which is a common issue for the count data, it will not be appropriate to be modeled by Poisson any more. In this case, the negative binomial (NB) distribution is much more widely used in modeling over-dispersed count data. In addition, the excess zeros is another common issue for count data, which also presents in our dataset by taking look at the histogram plot and descriptive table for multiplicity in chapter 3. Therefore, zero-inflated count models, including zero-inflated Poisson model and zero-inflated negative binomial model were fitted to balance excess zeros.

Chapter 2: Experiment design and methods

2.1 Animals

Animal care and use were conducted according to established guidelines approved by the Institutional Animal Care and Use Committee at the University of Alabama at Birmingham. Animals were treated humanely and with regard for alleviation of suffering. All animals were housed in a temperature controlled facility with a 12 hour light/dark cycle. Female Sprague Dawley CD rats (Charles River, Raleigh, NC) were bred and observed for the presence of sperm. Sperm positive females were separated, housed in polypropylene cages with glass water bottles (both polycarbonate/BPA free), and fed the phytoestrogen-free AIN-93G pelleted diet.

2.2 Treatments, diets and groups

Beginning on postnatal day two and continuing through postnatal day 20, the lactating dam of each litter was intragastrically gavaged with four different treatment, BPA, DEHP, LC(low combination of BPA and DEHP), HC (high combination of BPA and DEHP) and control

treatment SO (sesame oil) per day (on Mondays – Fridays only) along with diet (with/without genistein). There were 33 female offspring in each group (BPA, DEHP, LC, HC, SO, SO+GEN, BPA+GEN, DEHP+GEN, LC+GEN, and LC+GEN) respectively, all derived from individual litters, except one missing value for each of BPA+GEN and SO+GEN group. Offspring were palpated twice weekly to monitor tumor development. Data were recorded on palpable tumor latency, location, tumor burden, and multiplicity. Animals underwent necropsy when tumor burden exceeded 10% of body weight.

2.3 DMBA induced model

At 50 days postpartum, one female offspring from each litter of each treatment group was given a single gavage of 30 mg DMBA/kg BW. This dose results in a low number of mammary tumors and allows chemicals that predispose for mammary cancer to increase the number of mammary adenocarcinomas [66]. Because dam treatment for lactational exposure results in a one exposure compartment, only one offspring from each litter was used in each experiment.

2.4 Statistical methods

The time-to-event data, e.g., time-to-first-tumor (latency) and time-to-sacrifice (tumor burden), were analyzed using the LIFETEST, PHREG and LIFEREG procedures in SAS (SAS Institute Inc., Cary, NC). Survivor functions were first estimated for each group using the Kaplan-Meier method and compared across the ten groups using the Wilcoxon log-rank test and parametrically using survival regression analysis [12]. Those animals that had not developed a tumor by the end of the study or were sacrificed were treated as censored, and the end of study or sacrifice times were treated as censoring times. Values for latency and burden were expressed as median and mean \pm standard error of the mean.

Tumor multiplicity data were analyzed with the GENMOD (generalized linear models) and COUNTREG (count regression) procedure in SAS using Poisson regression, negative binomial regression, zero-inflated Poisson regression, and zero-inflated negative binomial regression on the tumor multiplicity [41]. Since there was a positive correlation between number of tumors and number of days in the study, the tests on multiplicity were done after adjusting for the number of days each animal was on the study. Values for multiplicity were presented as mean \pm standard error of the mean.

Chapter 3: Results

The total sample size for the initial study was 330 rats randomly assigned to 10 groups (BPA, DEHP, LC, HC, SO, SO+GEN, BPA+GEN, DEHP+GEN, LC+GEN, and HC+GEN), and each group had 33 rats. Two rats (one in the BPA+GEN and one in SO+GEN respectively) did not enroll into the study successfully due to unexpected reasons. These two missing value were deleted before the data analysis. As a result, the total sample size was 329.

3.1 Tumor latency

This section presented selective modeling and analysis results related to tumor latency.

3.1.1 General description

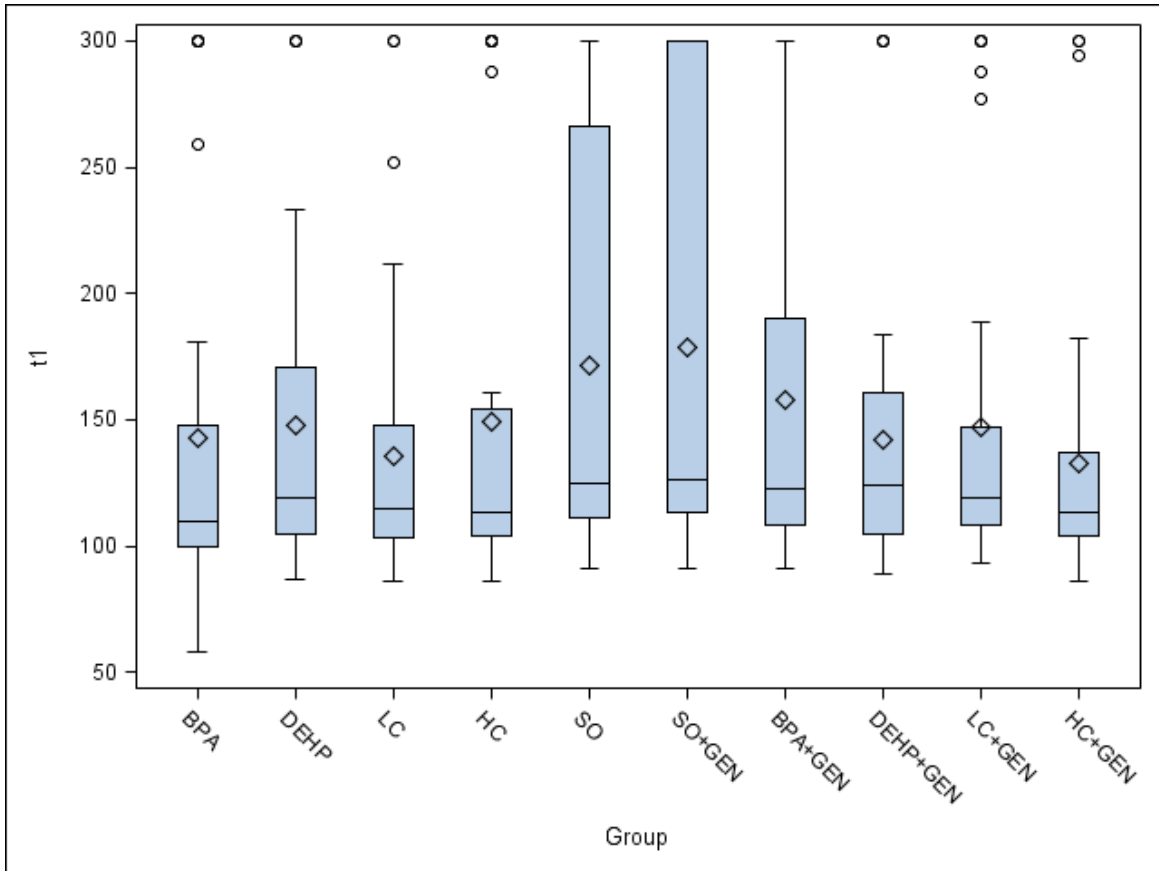
The descriptive statistics such as censor number and percentage, median of latency, mean \pm standard error, minimum/maximum by ten groups were shown in table 1 as followed. These statistics that grouped by five treatments was also presented in Figure 11 appendix 1. Obviously, there were quit large percentages for the censored data in the two control groups, SO and SO+GEN. In other words, large percentages of rats in the SO and SO+GEN group did not develop any tumor until the end of the study. The group that has shortest median latency is BPA

(110 days) followed by HC and HC+GEN (both 113 days), LC (115 days), DEHP and LC+GEN (both 119 days), BPA+GEN (112.5 days), and DEHP+GEN (124 days). The longest median latency group is the control group SO and SO+GEN, which had 125 and 126 days respectively. Diet with genistein tended to prolong the tumor latency from general, only except with the HC group. The boxplot in figure 1 gave more visual details. The median was much more stable than the mean. There was no significant difference for median while the means of two control groups were much larger than the treated and dieted groups. Especially, the means of the groups in which rats that were fed genistein as diet did not show an increasing pattern than the treatment groups. Therefore, we expected that the genistein did not affect the tumor latency in this study.

Table 1: Descriptive Table for Tumor Latency Sorted by Median

| Diet | Treatment | N | Censor/ Total (Rate) | Median | Mean(+/-stderr) | Min/Max |
|-------------|------------------|----------|-----------------------------|---------------|------------------------|----------------|
| Con | BPA | 33 | 4/29(12.1%) | 110.0 | 143 (+/- 11.99) | 58/300 |
| Con | HC | 33 | 5/28(15.2%) | 113.0 | 149 (+/- 12.9) | 86/300 |
| Gen | HC | 33 | 2/31(6.1%) | 113.0 | 132.91 (+/- 9.84) | 86/300 |
| Con | LC | 33 | 2/31(6.1%) | 115.0 | 135.61 (+/- 9.64) | 86/300 |
| Con | DEHP | 33 | 3/30(9.1%) | 119.0 | 147.94 (+/- 10.8) | 87/300 |
| Gen | LC | 33 | 3/30(9.1%) | 119.0 | 147.39 (+/- 11.6) | 93/300 |
| Gen | BPA | 32 | 4/28(12.5%) | 122.5 | 157.72 (+/- 12.41) | 91/300 |
| Gen | DEHP | 33 | 3/30(9.1%) | 124.0 | 142.36 (+/- 9.97) | 89/300 |
| Con | SO | 33 | 7/26(21.2%) | 125.0 | 171.48 (+/- 14.47) | 91/300 |
| Gen | SO | 32 | 10/22(31.3%) | 126.0 | 178.66 (+/- 15.1) | 91/300 |

Figure 1: Boxplot for Tumor Latency by Ten Groups



3.1.2 Survival analysis

In this subsection, the event for detecting first tumor was the target event and the survivor function was used to estimate the probability of this event occurs during the 300 cutoff days. Therefore, the data will be treated as censored data if the rat had never developed a tumor during the 300 study days.

The survivor functions were estimated by Kaplan-Meier curves for each group as presented in figure 2. The medians were exactly the same with descriptive table since there was no censored data before 300 days. And the differences of survivor function across ten group were compared by log-rank test, which was not significant with $P=0.1413$. However, the differences

across five treatment was significant with $p = 0.034$ by adjusting the diet. The P value for testing diet was 0.6471, which indicated that genistein did not delay the time to first tumor. So even though the two control groups (SO and SO+GEN) seems had better performance at tumor latency by comparing with other groups, the differences were not statistically significant. According to the results for strata five treatments by adjusting diet, the survival regression might be needed to include treatment and diet together into the model instead of using nonparametric way to estimate ten groups.

Figure 2: Kaplan-Meier Curves for Tumor Latency by Ten Groups

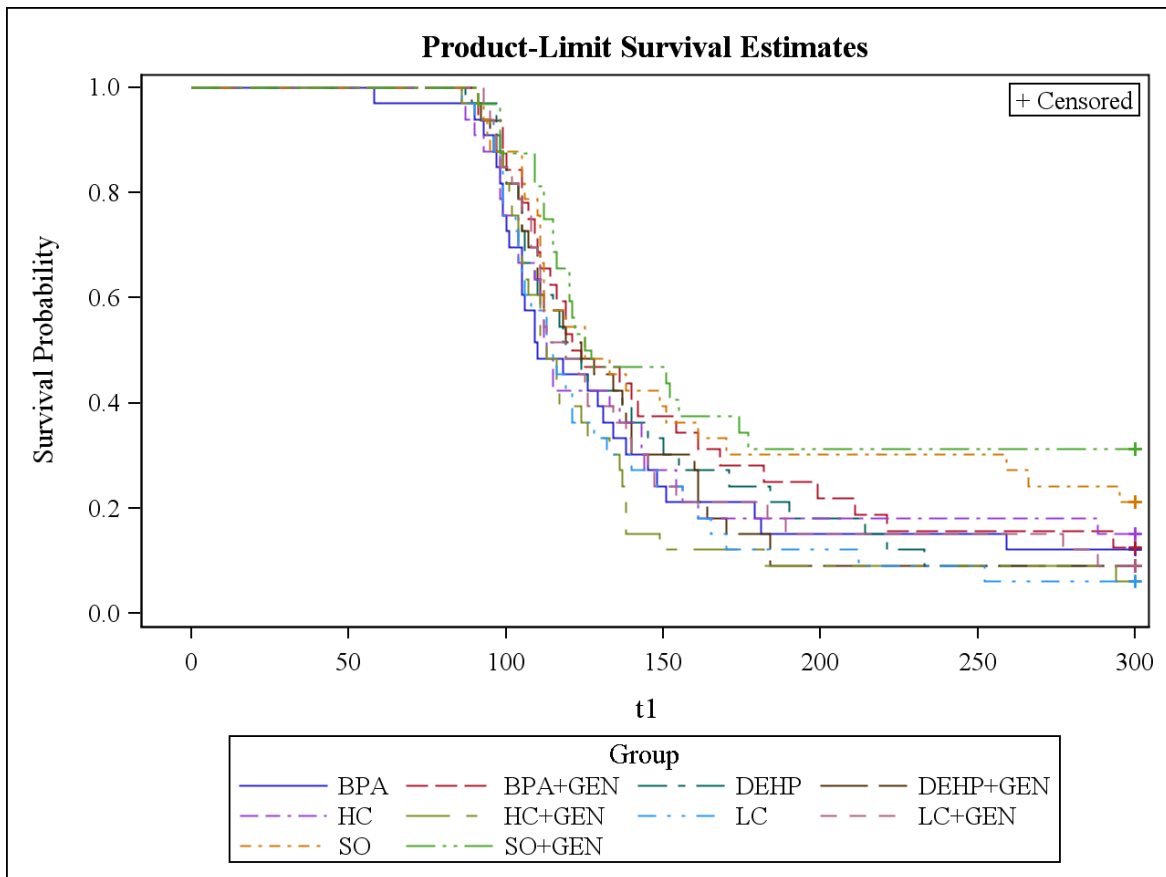


Table 2: Estimates for Tumor Latency Sored by Median from LifeTest procedure

| Diet | Treatment | N | Censor | MedianLT (95% CI) | MeanLT (+/-stderr) |
|-------------|------------------|----------|---------------|--------------------------|---------------------------|
| Con | BPA | 33 | 4/33(12.1%) | 110(105,134) | 138(10) |
| Con | HC | 33 | 5/33(15.2%) | 113(104,143) | 147.2(12.3) |
| Gen | HC | 33 | 2/33(6.1%) | 113(105,126) | 132.5(9.7) |
| Con | LC | 33 | 2/33(6.1%) | 115(104,128) | 132.7(8.2) |
| Con | DEHP | 33 | 3/33(9.1%) | 119(106,145) | 141.8(8.4) |
| Gen | LC | 33 | 3/33(9.1%) | 119(110,138) | 146.3(11.2) |
| Gen | BPA | 32 | 4/32(12.5%) | 122.5(110,161) | 156.8(12.1) |
| Gen | DEHP | 33 | 3/33(9.1%) | 124(110,140) | 131.8(5.5) |
| Con | SO | 33 | 7/33(21.2%) | 125(111,161) | 170.4(14.2) |
| Gen | SO | 32 | 10/32(31.3%) | 126(115,177) | 140.2(5.7) |

3.1.3 Model fitting

As shown in the figure 3, the negative log of survivor function plots were not linear, which suggested that the distribution for tumor latency was not exponential. Additionally, in figure 4, the log of negative log of survivor functions plots were not linear either, which means the Weibull model might not be appropriate either. Also, the log (-log (survivor)) plots were not approximately parallel, therefore that the proportional hazards model might not be appropriate to fit.

Figure 3: Negative log of Survivor function Plots for Tumor Latency by Ten Groups

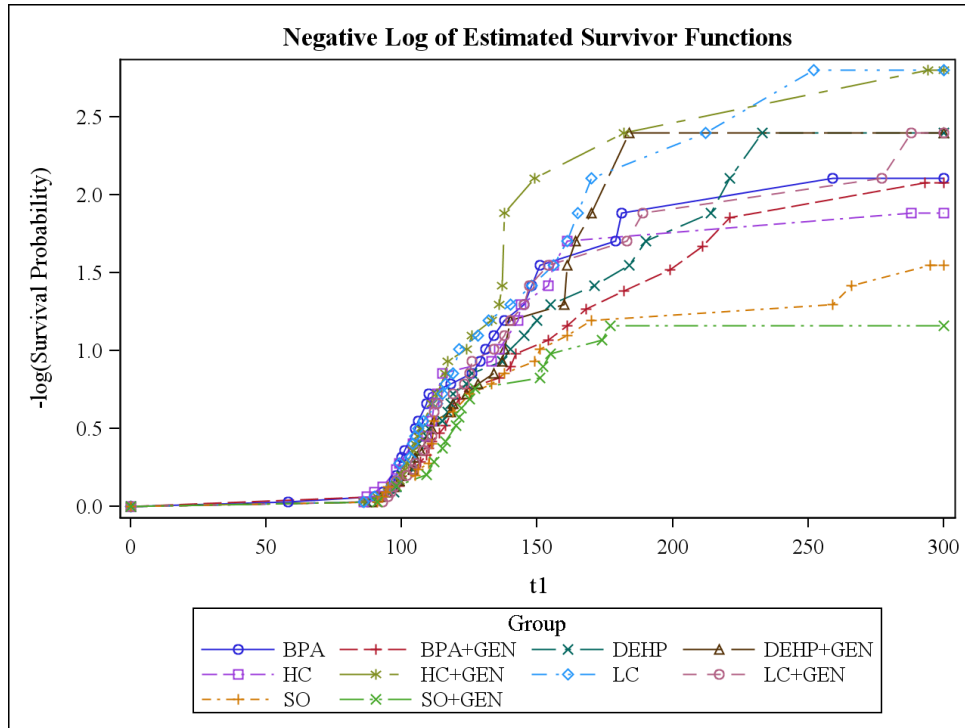
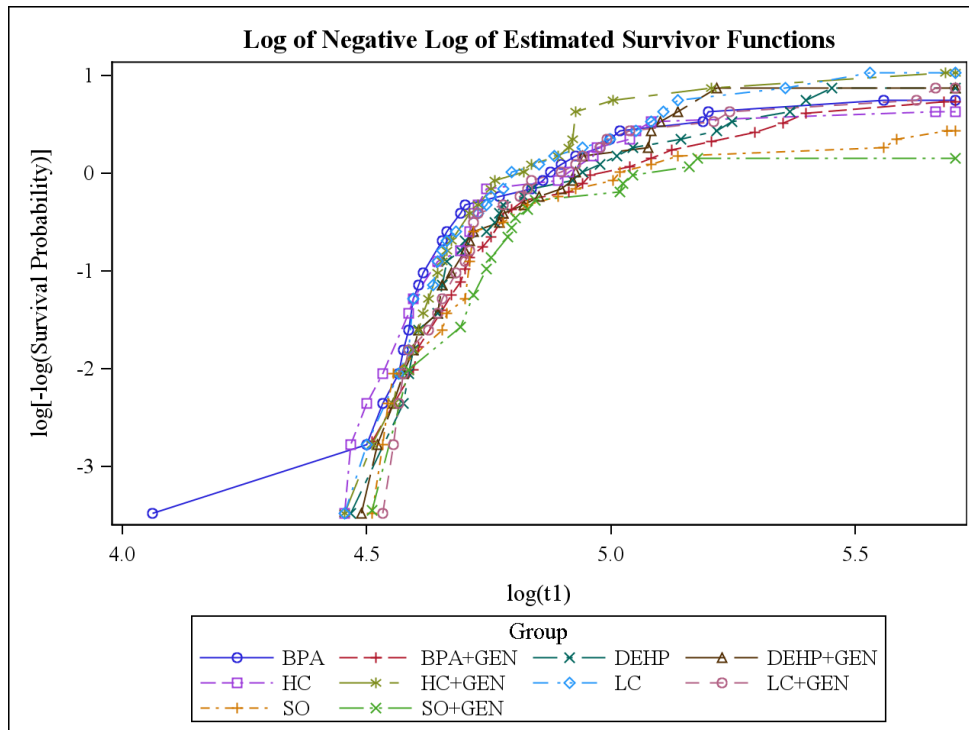


Figure 4: Log of Negative Log of Survivor Functions Plots for Tumor Latency by Ten Groups



First, the Proportional Hazards Model was fitted and model selection was performed by using PHREG procedure in SAS. Table 3 presented the details about the estimates of proportional hazards model for tumor latency. In the proportional hazards model, the effects of the covariates are to act multiplicatively on the hazard of the survival time, and therefore it is a little easier to interpret the corresponding hazard ratios than the regression parameters. The results shown that the hazard ratio of BPA, DEHP, LC, and HC were significantly larger than control group SO. Therefore, the treatment groups with different compounds could decrease the time for developing first tumor, but genistein and all the interactions did not shown any significant effect to the model. More specifically, the HC group had the largest hazard ratio than others, which means the High combination treatment decreased the tumor latency mostly. It was consistent with the information contained in the descriptive table, where both HC and HC+GEN had very small tumor latency.

Table 3: Estimates of Final Proportional Hazards Model for Tumor Latency

| Parameter | DF | Parameter Estimate | Standard Error | Chi-Square | Pr > ChiSq | Hazard Ratio |
|------------------|-----------|---------------------------|-----------------------|-------------------|----------------------|---------------------|
| BPA | 1 | 0.38433 | 0.19620 | 3.8372 | 0.0501 | 1.469 |
| DEHP | 1 | 0.42078 | 0.19428 | 4.6910 | 0.0303 | 1.523 |
| HC | 1 | 0.55019 | 0.19505 | 7.9569 | 0.0048 | 1.734 |
| LC | 1 | 0.53093 | 0.19364 | 7.5174 | 0.0061 | 1.701 |

However, the proportionality testing result ($P < .0001$) indicated that the proportional hazards assumption was not appropriate, same conclusion as the log of negative log of survivor functions plots in figure 4 indicated.

Secondly, a Weibull model was fitted by LIFEREG procedure in SAS and the parameter estimates for the final model were given in table 4. The results, similar with proportional hazards model, shown that all the treatments, including BPA, DEHP, LC, and HC significantly decreased the time for developing first tumor, but genistein and all the interactions did not shown any significant contribution to the model as well. In weibull model, the shortest tumor latency was LC group instead of HC group. At least these results indicated that the combinations of the BPA and DEHP had stronger carcinogenicity than single one.

Table 4: Estimates of Final Weibull Model for Tumor Latency

| Parameter | DF | Estimate | Std. Error | 95% Confidence Limits | | Chi-Square | Pr > ChiSq |
|----------------------|-----------|-----------------|-------------------|------------------------------|---------|-------------------|----------------------|
| Intercept | 1 | 5.4197 | 0.0694 | 5.2836 | 5.5557 | 6096.60 | <.0001 |
| BPA | 1 | -0.2413 | 0.0942 | -0.4260 | -0.0566 | 6.55 | 0.0105 |
| DEHP | 1 | -0.3116 | 0.0936 | -0.4950 | -0.1282 | 11.09 | 0.0009 |
| HC | 1 | -0.3135 | 0.0931 | -0.4961 | -0.1310 | 11.33 | 0.0008 |
| LC | 1 | -0.3381 | 0.0928 | -0.5201 | -0.1561 | 13.26 | 0.0003 |
| SO | 0 | 0.0000 | . | . | . | . | . |
| Scale | 1 | 0.4807 | 0.0213 | 0.4407 | 0.5243 | | |
| Weibull Shape | 1 | 2.0805 | 0.0922 | 1.9074 | 2.2692 | | |

Table 5 presented the details for model selection. All the criteria which included AIC, BIC, and likelihood ratio test suggested only variable treatment should be included in the final model. Therefore, the final model for tumor latency could be written as

$$\text{Log (Tumor latency)} = \begin{cases} 5.4197, & \text{if treatment} = \text{SO} \\ 5.1784, & \text{if treatment} = \text{BPA} \\ 5.1062, & \text{if treatment} = \text{DEHP} \\ 5.1081, & \text{if treatment} = \text{HC} \\ 5.0816, & \text{if treatment} = \text{LC} \end{cases}$$

The expected tumor latency for each treatment in Weibull model was tended to larger than those of descriptive statistics since Weibull modeling was based on mean which was not as robust as median. Therefore, the prediction from weibull model might be poor. Besides, the fitness for Weibull might be not well since the log negative log of survivor function plot was not linear and the proportional hazards assumption did not hold. Therefore, another distribution might need to apply in the future research.

Table 5: Model Selection for Weibull Distribution for Tumor Latency

| Model | Variables | DF | Log Likelihood | AIC | BIC | Comparison | Likelihood ratio test | Critical Value |
|-------|---------------------------------|----|----------------|--------|--------|------------|-----------------------|----------------|
| 0 | Intercept only | 2 | -295.926 | 595.85 | 603.44 | -- | -- | -- |
| 1 | Treatment | 6 | -286.538 | 585.08 | 607.83 | -0 vs 1 | 18.776* | 9.49 |
| 2 | Treatment, Diet | 7 | -286.521 | 587.04 | 613.59 | 1 vs 2 | 0.0331 | 3.84 |
| 3 | Treatment, Diet, Treatment*Diet | 11 | -284.353 | 590.71 | 632.43 | 2 vs 3 | 4.3373 | 9.49 |

Note: * P < .05.

3.2 Tumor burden

This section presented selective modeling and analysis results related to tumor burden.

3.2.1 General description

The descriptive statistics such as censor number and percentage, median of burden, mean +/- standard error, minimum/maximum were shown in table 6 as followed . These statistics that grouped by five treatments was also presented in Figure 12 appendix 1. Compared to tumor

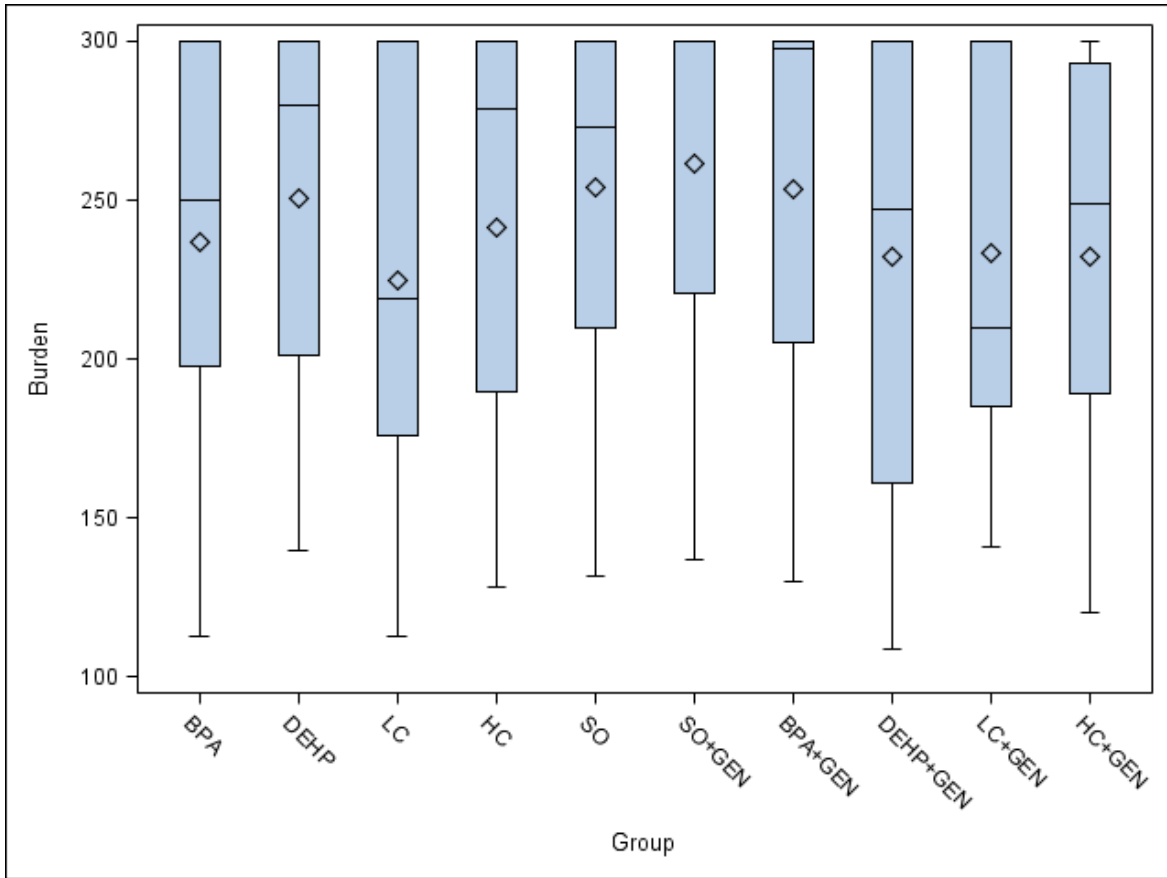
latency, there were quit large percentages for the censored data in each group, which means there were quit large percentage mice had not reached the criteria to sacrifice after 300 days. It raised the concern for the rate of the tumor development. The DMBA induced model might not succeed since some of rats had never developed a tumor while some of the rats did not have a large tumor growth rate even it had tumor onset.

The group that has shortest median burden is LC+GEN (210 days) followed by LC (219 days) and DEHP+GEN (247 days), HC+GEN (249 days), BPA (250 days), SO (273 days), HC (279 days), and DEHP (280 days) . The longest median burden group was the SO+GEN (300 days), which indicated that more than half of the rats in the SO+GEN group can survive beyond 300 days. As shown in the boxplot (figure 5), only LC group tended to shrink the tumor burden compared with other groups.

Table 6: Descriptive Table for Tumor Burden Sored by Median

| Diet | Treatment | N | Censor for Burden | Median | Mean(+/-stderr) | Min/Max |
|-------------|------------------|----------|--------------------------|---------------|------------------------|----------------|
| Gen | LC | 33 | 10/23(30.3%) | 210.0 | 233.48 (+/- 9.89) | 141/300 |
| Con | LC | 33 | 11/22(33.3%) | 219.0 | 224.61 (+/- 11.67) | 113/300 |
| Gen | DEHP | 33 | 12/21(36.4%) | 247.0 | 232.03 (+/- 12.22) | 109/300 |
| Gen | HC | 33 | 8/25(24.2%) | 249.0 | 232.15 (+/- 10.65) | 120/300 |
| Con | BPA | 33 | 10/23(30.3%) | 250.0 | 236.76 (+/- 10.48) | 113/300 |
| Con | SO | 33 | 12/21(36.4%) | 273.0 | 254.27 (+/- 9.36) | 132/300 |
| Con | HC | 33 | 15/18(45.5%) | 279.0 | 241.3 (+/- 11.56) | 128/300 |
| Con | DEHP | 33 | 16/17(48.5%) | 280.0 | 250.88 (+/- 9.57) | 140/300 |
| Gen | BPA | 32 | 16/16(50%) | 297.5 | 253.22 (+/- 11.12) | 130/300 |
| Gen | SO | 32 | 17/15(53.1%) | 300.0 | 261.44 (+/- 10.25) | 137/300 |

Figure 5: Boxplot for Tumor Burden by Ten Groups



3.2.2 Survival analysis

In this subsection, the event for sacrificing the rats was the target event and the survivor function was used to estimate the probability of the rats that could survive more than 300 days. Therefore, the data will be treated as censored data if the rat had not been sacrificed after the endpoint of the study.

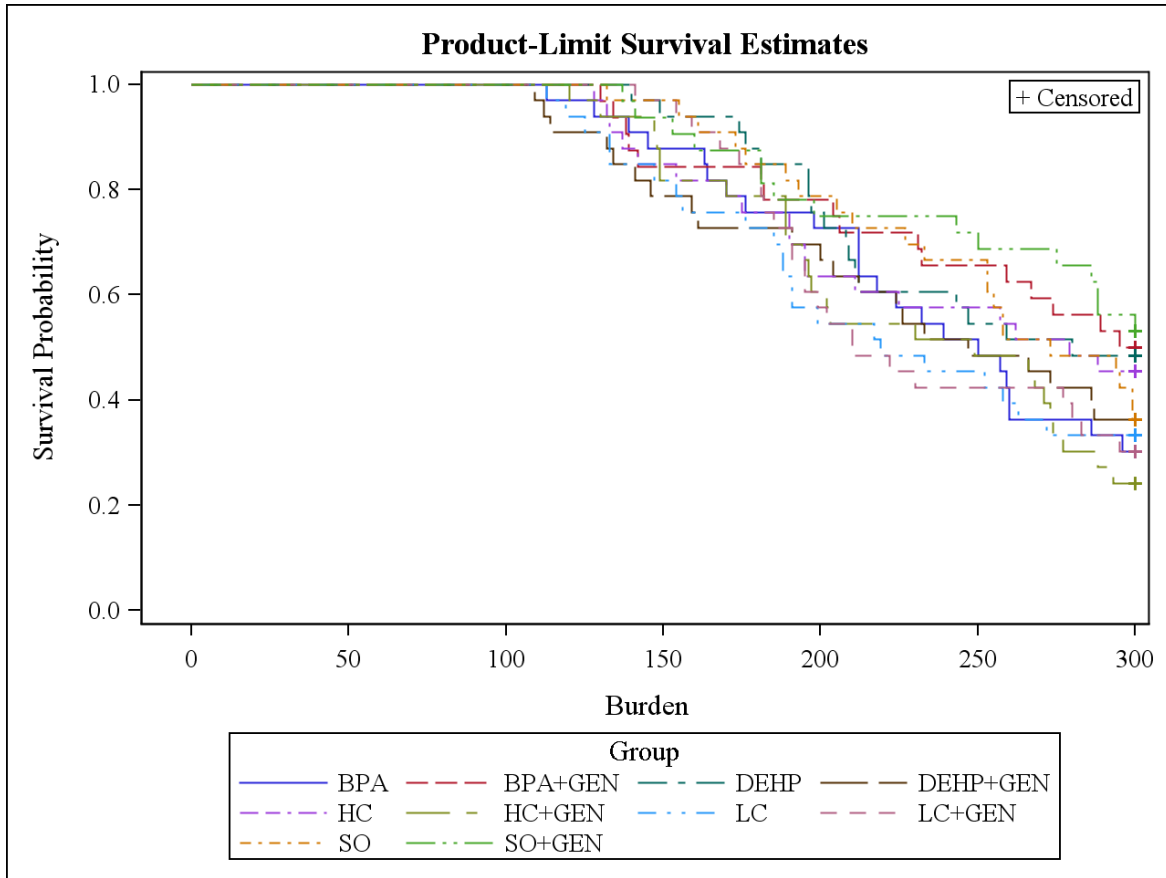
The survivor functions were estimated by Kaplan-Meier curves for each group as presented in figure 6. And the differences of survivor function across ten groups and five treatment by adjusting diet were compared by log-rank test, which both were not significant ($P=0.2273$ for ten groups, $P=0.2642$ for five treatments). Obviously, there were large amounts of

censored data in each group, which was also reflected in the descriptive table (table 6) and estimated table (table 7). Note that the mean survival time and its standard error were underestimated because the largest observation was censored and the estimation was restricted to the largest event time. And the median and upper limit of BPA+GEN and SO+GEN was indeterminate and represented as missing value since there were more than 50% censored data in these two groups. Therefore, all the information we knew about the median tumor burden of these two groups by nonparametric methods were they were more than 300 days.

Table 7: Estimates for Tumor Burden by Each Group from LifeTest

| Diet | Treatment | N | Censor | Median(95% CI) | Mean (+/-Stderr) |
|-------------|------------------|----------|---------------|-----------------------|-------------------------|
| Con | BPA | 33 | 10/33(30.3%) | 250(212,286) | 235.5(10.3) |
| Gen | BPA | 32 | 16/32(50%) | .(231,.) | 250.7(11) |
| Con | DEHP | 33 | 16/33(48.5%) | 280(209,.) | 241.2(8.2) |
| Gen | DEHP | 33 | 12/33(36.4%) | 247(200,.) | 227.3(11.5) |
| Con | HC | 33 | 15/33(45.5%) | 279(195,.) | 235.8(10.9) |
| Gen | HC | 33 | 8/33(24.2%) | 249(194,274) | 230.5(10.4) |
| Con | LC | 33 | 11/33(33.3%) | 219(188,272) | 215.3(10) |
| Gen | LC | 33 | 10/33(30.3%) | 210(191,283) | 232(9.6) |
| Con | SO | 33 | 12/33(36.4%) | 273(233,.) | 253.9(9.4) |
| Gen | SO | 32 | 17/32(53.1%) | .(250,.) | 261.4(10.4) |

Figure 6: Kaplan-Meier Curves for Tumor Burden by Ten Groups



3.2.3 Model fitting

The nonparametric method could not estimate the tumor burden explicitly due to a large percentage of the censoring. Therefore the parametric modeling was applied to extrapolate the median of the ten groups. Firstly, a general idea about the distribution of the data can be obtained by reading the negative log of survivor function plots and the log of negative log of survivor functions plots shown in the figure 7 and 8. Obviously, there were a bunch of non-linear lines in the negative log of survivor function plots, which suggested that the distribution for tumor burden was not exponential. In figure 8, however, the log of negative log of survivor functions

plots were approximately parallel linearly distributed, which means the proportional hazards model and the Weibull model might be appropriate.

Figure 7: Negative log of Survivor function for Tumor Burden by Ten Groups

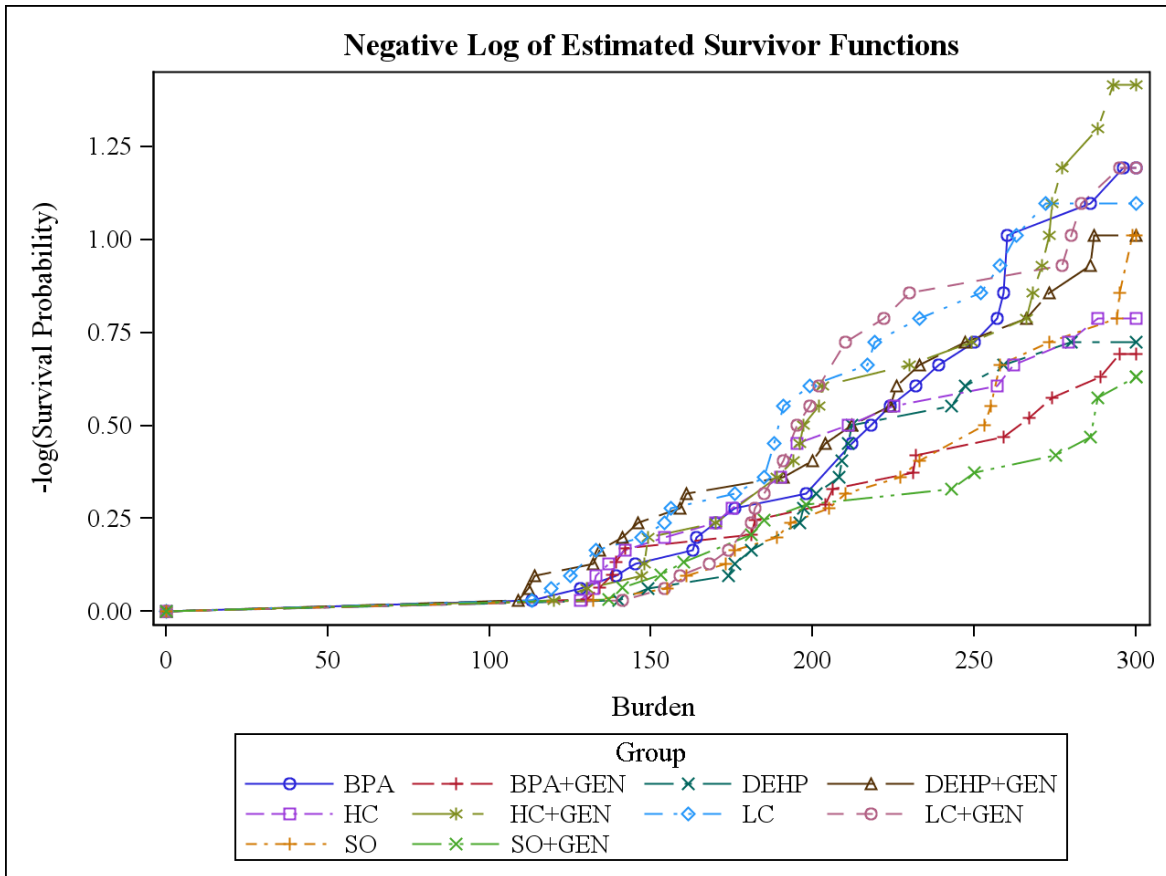
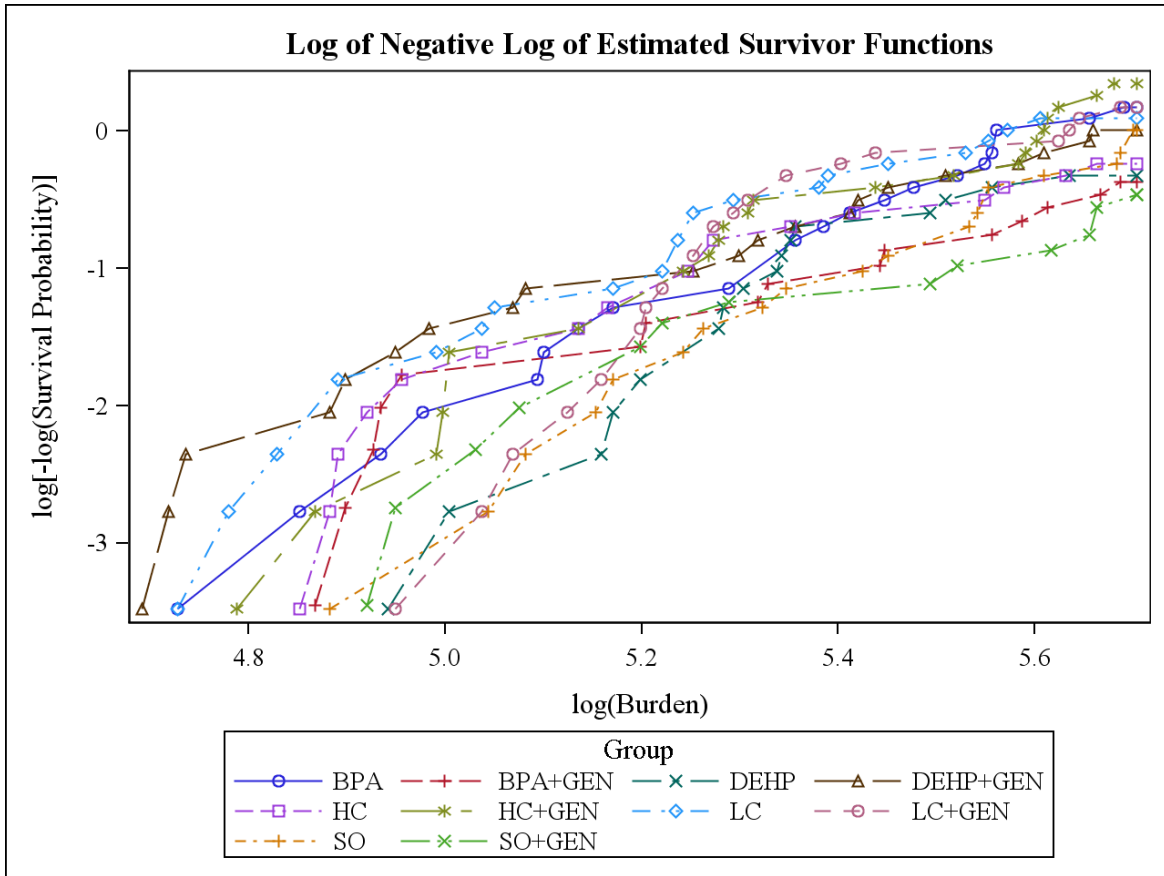


Figure 8: Log of Negative Log of Survivor Functions for Tumor Burden by Ten Groups



Similar to tumor latency, the proportional hazards model for tumor burden was fitted and model selection was performed in the first step by using PHREG procedure in SAS. No matter which model selection methods had been used, there were no explanatory variables stayed into the model at the .05 significance level. Additionally, the proportionality testing result ($P < .0001$) indicated that the proportional hazards assumption was not appropriate. Therefore the Weibull model was fitted in the second step.

The estimates for Weibull model was shown in table 8 and the model selection step was given in table 9. Only treatment stayed in the model among the three categorical variables (treatment, diet and the two-factor interactions) after two steps of backward selection. However,

Type III Analysis of Effects test gave a large P value ($P = 0.2246$) which indicated that the overall treatment was not significant in the model, even though the LC group significantly decreased the tumor burden for the rats compared with other groups. Therefore, dummy variables for each treatment group were created and fitted in the Weibull model for further analyzing. After step by step backward variable elimination, the last variable LC along with .0654 P value, can only stay in the model at 0.1 significance level. Hence, the Weibull model was not significant for tumor burden either. More details about model selection were presented by table 9.

Table 8: Estimates of Weibull Model for Tumor Burden

| Parameter | DF | Estimate | Std. Error | 95% Confidence Limits | | Chi-Square | Pr > ChiSq |
|----------------------|-----------|-----------------|-------------------|------------------------------|---------|-------------------|----------------------|
| Intercept | 1 | 5.7843 | 0.0544 | 5.6776 | 5.8910 | 11288.6 | <.0001 |
| BPA | 1 | -0.0627 | 0.0745 | -0.2088 | 0.0833 | 0.71 | 0.3998 |
| DEHP | 1 | -0.0587 | 0.0750 | -0.2057 | 0.0883 | 0.61 | 0.4339 |
| HC | 1 | -0.1147 | 0.0730 | -0.2577 | 0.0284 | 2.47 | 0.1161 |
| LC | 1 | -0.1599 | 0.0723 | -0.3017 | -0.0181 | 4.88 | 0.0271 |
| SO | 0 | 0.0000 | . | . | . | . | . |
| Scale | 1 | 0.3223 | 0.0200 | 0.2853 | 0.3641 | | |
| Weibull Shape | 1 | 3.1029 | 0.1930 | 2.7467 | 3.5052 | | |

Table 9: Model Selection for Weibull Distribution for Tumor Burden

| Model | Variables | DF | Log Likelihood | AIC | BIC | Comparison | Likelihood ratio test | P=.05 Critical Value |
|-------|---------------------------------|----|----------------|--------|--------|------------|-----------------------|----------------------|
| 0 | Intercept only | 2 | -232.039 | 468.08 | 475.66 | -- | -- | -- |
| 1 | LC | 3 | -230.434 | 466.87 | 478.25 | 0 vs 1 | 3.21 | 3.84 |
| 2 | Treatment | 6 | -229.185 | 470.37 | 493.13 | 1 vs 2 | 2.498 | 7.82 |
| 3 | Treatment, Diet | 7 | -229.182 | 472.36 | 498.92 | 2 vs 3 | 0.004 | 3.84 |
| 4 | Treatment, Diet, Treatment*Diet | 11 | -225.435 | 472.87 | 514.59 | 3 vs 4 | 7.494 | 9.49 |

3.3 Tumor multiplicity

The tumor multiplicity was analyzed by GENMOD procedure and COUNTREG procedure in SAS. Four different models, including Poisson, Negative Binomial, Zero-inflated Poisson, and Zero-inflated Negative Binomial, were fitted respectively. Since there was a positive correlation between number of tumors and number of days in the study, the tests on multiplicity were done after adjusting for the number of days each animal was on the study by set up the OFFSET option in the model.

3.3.1 General description

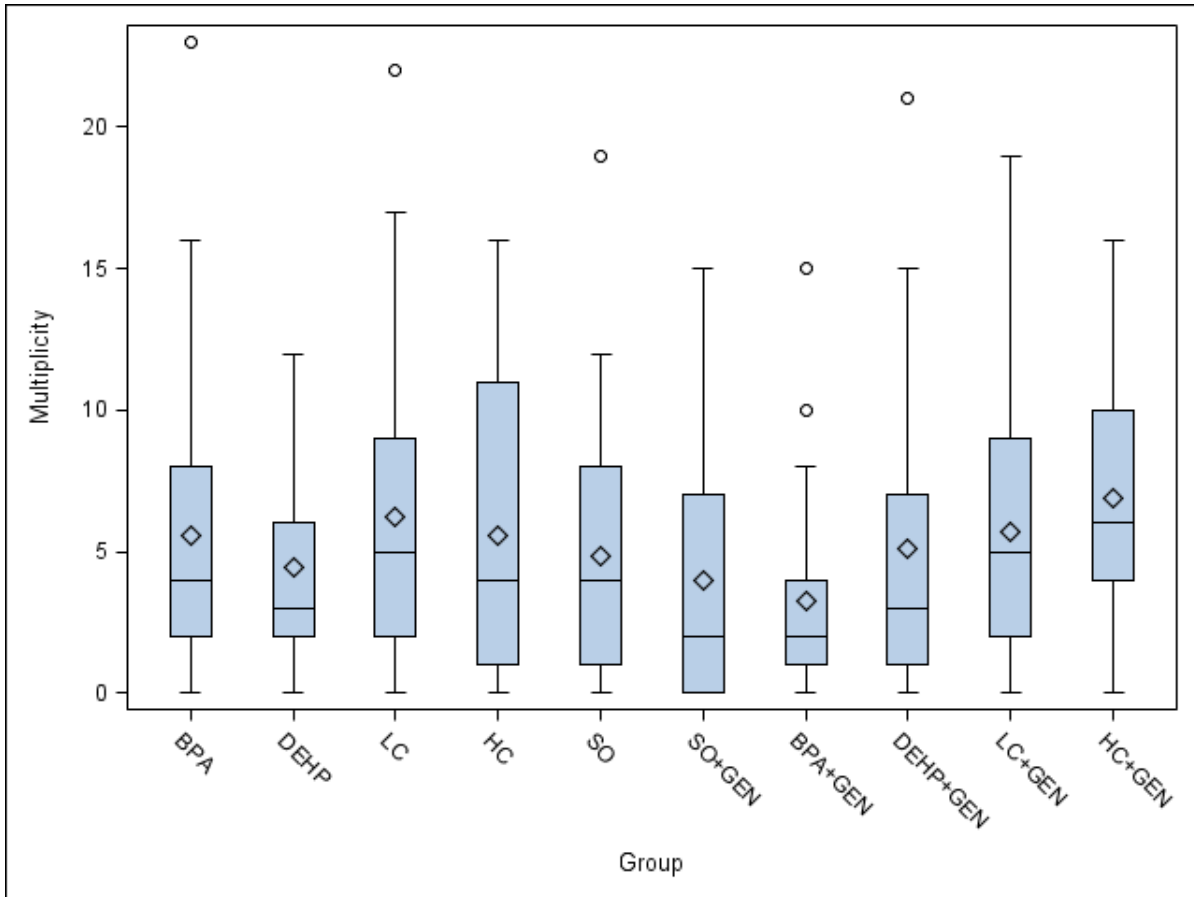
The descriptive statistics such as median of tumor multiplicity, mean +/- standard error, minimum/maximum, and median of tumor rate were shown in table 10 as followed. These statistics that grouped by five treatments was also presented in Figure 13 appendix 1. In general, the mean of tumor multiplicity was 5.15, with a larger variance 21.95, which raised over-dispersion problem. In addition, the zero fractions of BPA, HC, BPA+GEN and total exceeded 10%, especially the SO and SO+GEN group (21.21% and 31.25%, respectively), which

might imply that it is a mixture Poisson model with excess zeros. It might have heavy tail because of some extreme larger values. Specifically, the group that has smallest median multiplicity is SO+GEN and BPA +GEN (both 2) followed by DEHP and DEHP+GEN (both 3), HC, BPA, SO (all 4), LC and LC+GEN (both 5). The largest median multiplicity group was the HC+GEN (6), with 7.66% median of tumor rate. The boxplot with more details was presented in figure 9.

Table 10: Descriptive Table for Tumor Multiplicity Sored by Diet and Treatment

| Diet | Treatment | N | Median | Mean(+/-stderr) | Min/Max | Zero Fraction |
|-------------|------------------|----------|---------------|------------------------|----------------|----------------------|
| Gen | BPA | 32 | 2 | 3.22 (+/- 0.56) | 0/15 | 12.50 % |
| Gen | SO | 32 | 2 | 3.97 (+/- 0.77) | 0/15 | 31.25 % |
| Con | DEHP | 33 | 3 | 4.45 (+/- 0.61) | 0/12 | 9.09 % |
| Gen | DEHP | 33 | 3 | 5.09 (+/- 0.85) | 0/21 | 9.09 % |
| Con | SO | 33 | 4 | 4.82 (+/- 0.79) | 0/19 | 21.21 % |
| Con | BPA | 33 | 4 | 5.58 (+/- 0.93) | 0/23 | 12.12 % |
| Con | HC | 33 | 4 | 5.58 (+/- 0.92) | 0/16 | 15.15 % |
| Con | LC | 33 | 5 | 6.21 (+/- 0.92) | 0/22 | 6.06 % |
| Gen | LC | 33 | 5 | 5.73 (+/- 0.82) | 0/19 | 9.09 % |
| Gen | HC | 33 | 6 | 6.88 (+/- 0.8) | 0/16 | 6.06 % |
| Total | - | 328 | 4 | 5.15(+/- 4.68) | 0/23 | 13.11 % |

Figure 9: Boxplot for Tumor Multiplicity by Ten Groups



3.3.2 Poisson regression

The full Poisson regression model was fitted with treatment, diet, and two-factor interactions in the first step. However, the variable diet can be dropped by rearrange dummy variable of the interactions (Diet = Diet * Treatment SO). Poisson regression detected that treatment BPA and LC increased the multiplicity compared with other groups. The P value for DEHP was close to 0.5 but still not significant. Besides, interaction between diet and BPA decreased the tumor multiplicity while diet interacted with HC increased multiplicity, compared with other interactions. Table 11 presented the details of final Poisson model for tumor multiplicity. Likelihood ratio test for model selection were illustrated in **Error! Reference**

source not found. appendix 1, the full model in table 11 was the final model which is significantly explained the majority of the variability of tumor multiplicity.

Table 11: Final Poisson Model for Tumor Multiplicity

| Parameter | DF | Estimate | Standard Error | Wald 95% Confidence Limits | | Wald Chi-Square | Pr > ChiSq |
|------------------|-----------|-----------------|-----------------------|-----------------------------------|---------|------------------------|----------------------|
| Intercept | 1 | -3.9660 | 0.0793 | -4.1214 | -3.8106 | 2500.95 | <.0001 |
| BPA | 1 | 0.2174 | 0.1083 | 0.0052 | 0.4296 | 4.03 | 0.0447 |
| DEHP | 1 | -0.0650 | 0.1144 | -0.2893 | 0.1592 | 0.32 | 0.5698 |
| HC | 1 | 0.1984 | 0.1083 | -0.0138 | 0.4106 | 3.36 | 0.0669 |
| LC | 1 | 0.3782 | 0.1057 | 0.1710 | 0.5853 | 12.81 | 0.0003 |
| SO | 0 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| Diet*BPA | 1 | -0.6167 | 0.1231 | -0.8578 | -0.3755 | 25.11 | <.0001 |
| Diet*DEHP | 1 | 0.2116 | 0.1129 | -0.0097 | 0.4330 | 3.51 | 0.0609 |
| Diet*HC | 1 | 0.2487 | 0.0992 | 0.0543 | 0.4431 | 6.28 | 0.0122 |
| Diet*LC | 1 | -0.1200 | 0.1008 | -0.3177 | 0.0776 | 1.42 | 0.2339 |
| Diet*SO | 1 | -0.2217 | 0.1190 | -0.4550 | 0.0115 | 3.47 | 0.0624 |
| Scale | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | | |

3.3.3 Negative binomial regression

Negative Binomial was fitted due to the over-dispersion problem in Poisson regression. The over-dispersion parameter α is significantly different from zero in negative binomial regression. That means the Poisson regression might not be valid since the assumption of equal variance and mean was violated. Beside, only the interaction between diet and BPA significantly decreased tumor multiplicity compared with other interactions, which was consistent with the

result of Poisson model. Similarly, variable diet can be dropped and be represented as the level of interaction between diet and treatment SO in the full model. Table 12 gave the parameter estimates and Wald test for individual variables in the model. Variable selection was performed by likelihood ratio test for NB model in Table 17 appendix 1 as well. Nevertheless, the results indicated that none of the variables stayed in the final model.

Table 12: Negative Binomial Model for Tumor Multiplicity

| Parameter | DF | Estimate | Standard Error | Wald 95% Confidence Limits | | Wald Chi-Square | Pr > ChiSq |
|------------------|-----------|-----------------|-----------------------|-----------------------------------|---------|------------------------|----------------------|
| Intercept | 1 | -3.8064 | 0.1947 | -4.1881 | -3.4248 | 382.04 | <.0001 |
| BPA | 1 | 0.2181 | 0.2737 | -0.3182 | 0.7545 | 0.64 | 0.4254 |
| DEHP | 1 | -0.0828 | 0.2761 | -0.6241 | 0.4584 | 0.09 | 0.7642 |
| HC | 1 | 0.2862 | 0.2739 | -0.2506 | 0.8231 | 1.09 | 0.2960 |
| LC | 1 | 0.3962 | 0.2727 | -0.1383 | 0.9306 | 2.11 | 0.1463 |
| SO | 0 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| Diet*BPA | 1 | -0.6350 | 0.2822 | -1.1881 | -0.0819 | 5.06 | 0.0244 |
| Diet*DEHP | 1 | 0.2708 | 0.2757 | -0.2696 | 0.8112 | 0.96 | 0.3261 |
| Diet*HC | 1 | 0.1099 | 0.2702 | -0.4197 | 0.6394 | 0.17 | 0.6843 |
| Diet*LC | 1 | -0.1853 | 0.2705 | -0.7154 | 0.3448 | 0.47 | 0.4932 |
| Diet*SO | 1 | -0.2268 | 0.2801 | -0.7758 | 0.3223 | 0.66 | 0.4183 |
| Scale | 1 | 1.0295 | 0.0965 | 0.8568 | 1.2371 | | |

3.3.4 Zero-inflated Poisson regression

According to the descriptive table in table 10, there were a certain amount of rats that never developed any tumor in each group before the endpoint of the study. This means there

might be a mixture Poisson model with excess zeros for tumor multiplicity. The histogram plot for tumor multiplicity was presented in Figure 14 appendix 1. Although the overall zero fractions were not very high, there were quite high percentages in some of the individual groups, especially two control groups SO and SO+GEN. The histograms of tumor multiplicity for individual group were omitted in this thesis due to the limitation of pages.

In order to solve the issue of excess zeros in Poisson model, Zero-inflated Poisson model was introduced. Compared with control treatment group, all the levels of treatment had less zeros in the zero-inflated model, and were significantly different from each other. LC treatment group had smallest zero fraction followed by DEHP, HC and BPA. Meanwhile, the DEHP and interaction of Diet and Treatment BPA decreased tumor multiplicity significantly compared with other groups, in the Poisson model. Likelihood ratio test suggested that the treatment and the interaction of treatment and diet should be included into the final model (see Table 18 in appendix 1).

Table 13: ZIP Model for Tumor Multiplicity

| Parameter | DF | Estimate | Standard Error | Wald 95% Confidence Limits | | Wald Chi-Square | Pr > ChiSq |
|------------------|-----------|-----------------|-----------------------|-----------------------------------|---------|------------------------|----------------------|
| Intercept | 1 | 1.8091 | 0.0797 | 1.6529 | 1.9654 | 515.00 | <.0001 |
| BPA | 1 | 0.0366 | 0.1089 | -0.1768 | 0.2500 | 0.11 | 0.7368 |
| DEHP | 1 | -0.2275 | 0.1158 | -0.4544 | -0.0005 | 3.86 | 0.0495 |
| HC | 1 | 0.0715 | 0.1089 | -0.1420 | 0.2850 | 0.43 | 0.5116 |
| LC | 1 | 0.0788 | 0.1061 | -0.1292 | 0.2869 | 0.55 | 0.4576 |
| SO | 0 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| Diet*BPA | 1 | -0.5688 | 0.1267 | -0.8172 | -0.3204 | 20.14 | <.0001 |
| Diet*DEHP | 1 | 0.1373 | 0.1146 | -0.0873 | 0.3618 | 1.44 | 0.2309 |

| | | | | | | | |
|----------------------|---|---------|--------|---------|---------|-------|--------|
| Diet*HC | 1 | 0.1100 | 0.0996 | -0.0853 | 0.3053 | 1.22 | 0.2696 |
| Diet*LC | 1 | -0.0497 | 0.1014 | -0.2484 | 0.1490 | 0.24 | 0.6242 |
| Diet*SO | 1 | -0.0600 | 0.1201 | -0.2954 | 0.1753 | 0.25 | 0.6170 |
| Scale | 0 | 1.0000 | 0.0000 | 1.0000 | 1.0000 | | |
| Inf_Intercept | 1 | -1.0487 | 0.2845 | -1.6064 | -0.4911 | 13.59 | 0.0002 |
| Inf_BPA | 1 | -1.0315 | 0.5058 | -2.0228 | -0.0401 | 4.16 | 0.0414 |
| Inf_DEHP | 1 | -1.3188 | 0.5363 | -2.3700 | -0.2676 | 6.05 | 0.0139 |
| Inf_HC | 1 | -1.0943 | 0.4940 | -2.0626 | -0.1260 | 4.91 | 0.0268 |
| Inf_LC | 1 | -1.4748 | 0.5536 | -2.5598 | -0.3899 | 7.10 | 0.0077 |

3.3.5 Zero-inflated negative binomial regression

The last step for model fitting, zero-inflated negative binomial model was fitted to solve both the over-dispersion and excess zeros issues (see table 14). The over-dispersion parameter alpha is significantly different from zero in the main model, which supported our assumption about the over-dispersion. Beside, only interaction between diet and treatment BPA significantly decreased the tumor multiplicity compared with other interaction groups, which was consistent with previous results. But no significantly contribution was from main effect treatment. These results were very similar with those of negative binomial regression. Another similar result with NB regression was the variable selection in the main model. None of the three variables (treatment, diet and two-factor interactions) significantly explained the variability of tumor multiplicity (see Table 19 in appendix 1).

Compared the zero-inflated model with ZIP, however, there were giant standard errors for inflated-BPA, inflated-DEHP, and inflated-LC, which brought a further concern about the validity of the zero-inflated model.

Table 14: ZINB model for tumor multiplicity

| Parameter | DF | Estimate | Standard Error | Wald 95% Confidence Limits | | Wald Chi-Square | Pr > ChiSq |
|----------------------|-----------|-----------------|-----------------------|-----------------------------------|----------|------------------------|----------------------|
| Intercept | 1 | -3.6526 | 0.2044 | -4.0532 | -3.2521 | 319.41 | <.0001 |
| BPA | 1 | 0.0612 | 0.2739 | -0.4756 | 0.5980 | 0.05 | 0.8232 |
| DEHP | 1 | -0.2401 | 0.2764 | -0.7818 | 0.3015 | 0.75 | 0.3849 |
| HC | 1 | 0.1667 | 0.2829 | -0.3879 | 0.7213 | 0.35 | 0.5557 |
| LC | 1 | 0.2393 | 0.2729 | -0.2955 | 0.7742 | 0.77 | 0.3805 |
| SO | 0 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| Diet*BPA | 1 | -0.6368 | 0.2681 | -1.1623 | -0.1114 | 5.64 | 0.0175 |
| Diet*DEHP | 1 | 0.2701 | 0.2615 | -0.2425 | 0.7827 | 1.07 | 0.3018 |
| Diet*HC | 1 | 0.0900 | 0.2637 | -0.4268 | 0.6069 | 0.12 | 0.7328 |
| Diet*LC | 1 | -0.1843 | 0.2560 | -0.6861 | 0.3174 | 0.52 | 0.4715 |
| Diet*SO | 1 | -0.1481 | 0.2989 | -0.7340 | 0.4378 | 0.25 | 0.6203 |
| Dispersion | 1 | 0.9041 | 0.0985 | 0.7302 | 1.1193 | | |
| Inf_Intercept | 1 | -1.5951 | 0.4659 | -2.5082 | -0.6821 | 11.72 | 0.0006 |
| Inf_BPA | 1 | -20.8572 | 56196.83 | -110165 | 110122.9 | 0.00 | 0.9997 |
| Inf_DEHP | 1 | -22.8053 | 59530.33 | -116700 | 116654.5 | 0.00 | 0.9997 |
| Inf_HC | 1 | -2.0466 | 1.8719 | -5.7155 | 1.6223 | 1.20 | 0.2743 |
| Inf_LC | 1 | -22.5430 | 58282.82 | -114255 | 114209.7 | 0.00 | 0.9997 |

3.3.6 Model comparison

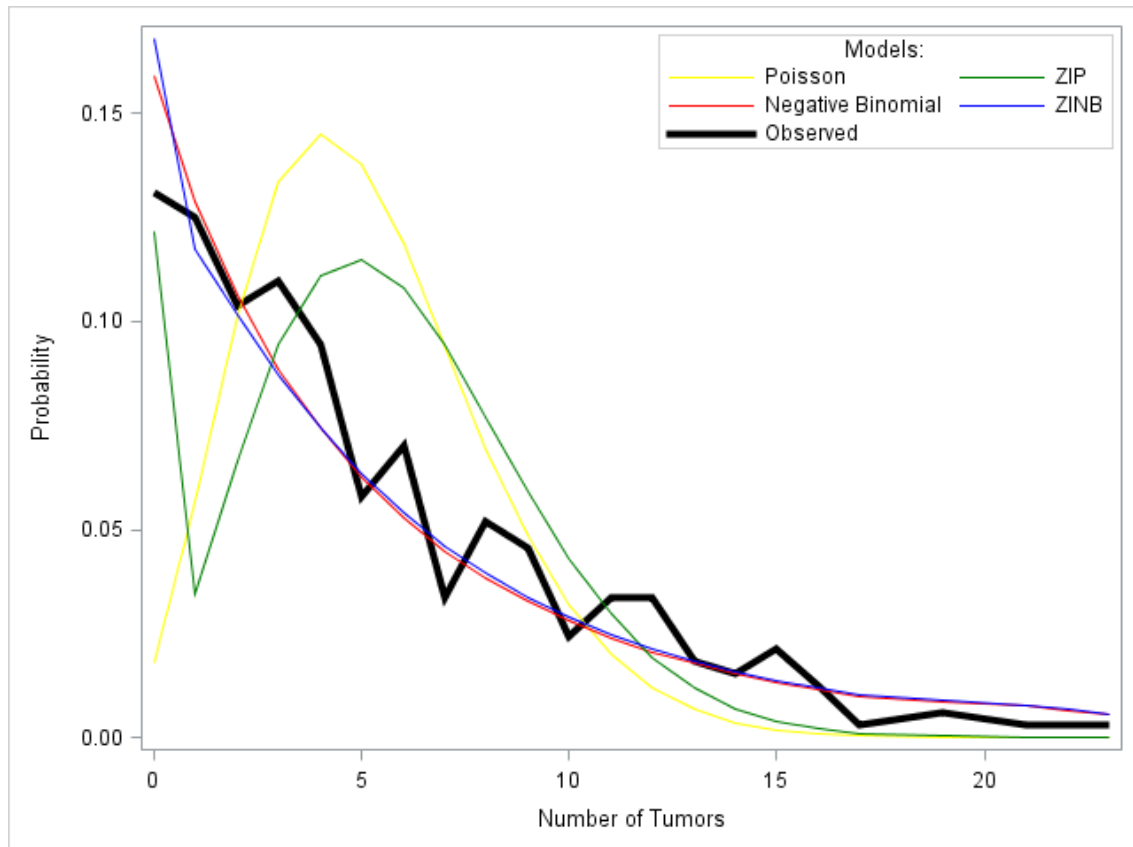
Due to the natural of the dataset, the four different regressions gave different results about the relationship between tumor multiplicity and treatment, diet, and interactions. Therefore, several criteria for model comparison, including log likelihood, AIC, BIC, Deviance/DF, Pearson Chi-Square/DF, and the estimated values when count equals zero were presented in table 18.

Generally speaking, NB and ZINB had better performance than Poisson and ZIP according to all the criteria in the table. Log-likelihood ratio test were performed between Poisson and ZIP, and between ZB and ZINB. And the results suggested that with respect to the detective power, ZIP was greater than Poisson, while no significantly difference between NB and ZINB. The Poisson gave a big improvement in log likelihood, AIC and BIC. Furthermore, ZIP model had priority on estimated value of count equals zero, considering 12.14% versus 13.11%. Pearson Chi-Square/DF indicated that NB had the best fitness for this data. Unfortunately, the global null hypotheses of NB and ZINB models were failed to reject. Same results could be extracted from the model comparisons plot in figure 10. Obviously, NB and ZINB had a better fitting performance, while ZIP had a better accuracy when the tumor number was zero, and a better explanation of our data. Poisson model had the worst fitting among the fours.

Table 15: Model Comparison for Four Full Models

| Models | Poisson | Negative binomial | ZIP | ZINB |
|---|-----------|-------------------|-----------|----------|
| DF | 9 | 10 | 13 | 14 |
| Log Likelihood | -1401.586 | -932.987 | -1037.162 | -929.234 |
| AIC | 2823.172 | 1887.973 | 2104.324 | 1890.468 |
| BIC | 2861.102 | 1929.696 | 2161.219 | 1951.156 |
| Deviance/DF | 5.809 | 1.177 | -- | -- |
| Pearson Chi-Square/DF | 6.954 | 1.042 | 2.476 | 1.095 |
| Estimated P ($\hat{Y}=0$) | 0.0178 | 0.1589 | 0.1214 | 0.1680 |
| / Observed P (y=0) | / 0.1311 | / 0.1311 | / 0.1311 | / 0.1311 |

Figure 10: Model Comparisons for Fitted and Observed Models of Tumor Multiplicity



Chapter 4: Discussion

4.1 Interpretation

This is a comprehensive experimental research which combined carcinogenesis study and chemoprevention study together. In general, the chemical compounds BPA, DEHP and their combinations, no matter low dosage or high dosage, are all carcinogens. On the other hand, the genistein did not show any chemopreventive effect statistically in this study. In order to interpret the results more robustly, the estimated median values were also presented in the thesis even though the procedures were all calculated based on the means, and the conclusion will be made by considering both results from the medians and the means.

First of all, there were some important results that been verified with respect to tumor latency. According to the Wilcox log-rank test, there is no significant difference of Kaplan-Meier curves across the ten groups (please refer to table 2 and figure 2). However, the significance had been detected among five treatments strata by adjusting diet in the same procedure (please refer to figure 11 in appendix 1). For further analysis, proportional hazards model and Weibull model were fitted. The same results shown that all the chemical treatment, including BPA, DEHP, Low combination and High combination, significantly shrank the time to first tumor which had been detected, meanwhile, they all significantly differ from each other as well (please refer to table 3 and 4). Among those chemical compounds, High combination group tended to have largest hazard ratio and shortest median value of tumor latency according to the result of proportional hazards model and the estimation of LIFETEST. However, in the Weibull model fitting procedure, Low combination group had the smallest estimated coefficient than others, which shortened the tumor latency mostly. For tumor latency, the proportional assumption was violated according to proportionality test in PHREG procedure. Therefore, the results from proportional

hazards model were not as reliable as the results from LIFETEST and LIFEREG procedure. The log likelihood ratio test was conducted to select the variables due to the large P values of some variables in the model. Finally, treatment with five levels and four dummy variables were stayed in the model and could explain the model sufficiently. As shown in table 2, the medians of tumor latency for all counterpart groups with genistein diet were tended to be larger than those of the groups without genistein diet, except for HC treatment. It means that the genistein tended to delay the time to first tumor but lack of statistical significance.

Additionally, compared the results of the tumor latency, these of tumor burden were not exactly consistent. Still, the nonparametric comparison of Kaplan-Meier curves and log-rank test across ten groups were not significant. Even the overall effect of treatment by adjusting diet was not significant either. And because there were two groups, BPA+GEN and SO+GEN, having more than half of censored data, the estimated medians were indeterminate by nonparametric and could only be extrapolated by parametric methods. Therefore, all the information we knew about the median tumor burden of these two groups by nonparametric methods were they were more than 300 days. As shown in table 7, there was a big gap between LC group and others, no matter with or without genistein diet. Same result could also be reflected from descriptive table and boxplot, which suggested that the Low combination treatment shrank the tumor burden a lot. In order to further analysis the treatment effect, parametric methods, such as proportional hazards regression and Weibull regression were performed. Similar to the tumor latency, the proportionality test for tumor burden had not been past, which indicated that the proportional assumption for tumor burden was violated. Besides, the model was not significant, and none of the variable could stay in the final model. So the Weibull model was conducted to obtain a better fitness. Although the log likelihood ratio test indicated that there was no significant difference

between the full model with all categorical variable treatment, diet and their interaction terms and the reduced model with only treatment, the treatment variable could not stay in the model at last. Therefore, neither of the two variables could explain the variability of tumor burden in this study, even though the Low combination group indeed shrank the tumor burden significantly by comparing with other treatments. It means if we only compare the difference of tumor burden across the five treatment groups, Low combination group had the shortest one. However, the overall treatment factor did not affect the tumor burden statistically. It might be due to the large percentage of censoring data. Also because the rats were sacrificed when the tumor volume exceeded 10% of body weight, the time to sacrifice was not as objective as time to death. Therefore, conclusions that derive from tumor burden were not appropriate to infer to the whole population.

Last but not least, tumor multiplicity had the most complicated analyzing results that needed to be discussed thoroughly. According to the descriptive table and boxplot, the smallest median and mean of tumor multiplicity was the BPA+GEN group, which was a big decrease compared to the BPA group. Similar to tumor latency, most of the medians and means of tumor multiplicity for counterpart groups with genistein diet were tended to be smaller than or equal to those of the groups without genistein diet. Nevertheless, the mean of DEHP, median and mean of HC treatment groups had the reverse pattern. It means the genistein might have effect on reducing total tumor number, and there might be some problems in DEHP and HC groups. In order to determine the explanatory effects of those chemical treatments and genistein diet on tumor multiplicity of the rats, general Poisson regression was conducted firstly since the count data was expected to follow a Poisson distribution[1]. The backward elimination suggested that the final model for Poisson regression should including treatment and the interaction between

treatment and diet. The main effect diet can be represented as the interaction level of treatment SO interacted with diet genistein. As shown in the table 11, the LC group had the largest coefficient among the main effect treatment, and HC interacted with genistein had the largest coefficient compared to other interactions. These were consistent with the information from descriptive table. The Poisson regression indicated that the BPA and Low combination groups significantly increased the total number of tumor per rat compared to other treatment effect, while the BPA interacted with genistein could reduce the tumor multiplicity. Also, the genistein combined with High combination group could increase the tumor amount. The main effect of genistein did not show up statistically in Poisson model.

However, the over-dispersion issue can be detected in the descriptive table (table 10), where the mean of tumor multiplicity was 5.15, with a larger variance 21.95. Therefore, the negative binomial regression was conducted as in the most common solution for over-dispersion. The parameter α for over-dispersion was significantly different from zero, which means the negative binomial model was more appropriate for this data than Poisson model. Yet after fitting this flexible model, the significance of the explanatory factors was gone. It had a similar situation of the Weibull model for tumor burden, with a significant interaction level of BPA * genistein, but none of the overall effect of the three predictors. It gave us a general idea of that the tumor multiplicity of BPA* genistein group was smaller than other interaction groups, which can be read from the descriptive table and consistent with the result of Poisson. However, the overall treatment, diet and their interaction would not affect the total number of tumor per rat. And the whole model was not statistically significant.

From the table 10 and figure 9, we can also observe a quite large fraction of zeros, especially in these two control groups, SO and SO+GEN. Zero-inflated count models,

corresponding to the Poisson and negative binomial, were appropriate for fitting this large fraction of zeros in count data according to the methodology introduction in chapter 1. Zero-inflated Poisson (ZIP) model was fitted first. After deducting some excess zeros in each group, the DEHP treatment and BPA * genistein showed significantly reducing effects on tumor multiplicity by comparing other groups. This was statistically reasonable because they had relatively small means and zero fractions. Therefore, when other groups were deflated after eliminating proportional zeros, the means of these groups will be increase so that the groups that had low means would have significantly lower means compared to other deflated groups. So the decreasing of tumor multiplicity could be detected by fitting ZIP model. However, biologically, this result was beyond our expectation. We did not believe genistein combined by BPA can have a better chemopreventive effect than genistein itself. There must some hidden biological reasons that had not been revealed through the experiment. On the other hand, the control group SO had largest zero fraction compare to all the other levels of treatment in the zero-inflated model, and they were significantly different from each other. In the last step, the corresponding zero-inflated negative binomial model was fitted to solve both the over-dispersion and excess zeros issues. Similar to negative binomial model, the over-dispersion parameter alpha was significant, which indicated that the ZIP model might not solve the over-dispersion problem. However, same to the NB model, only interaction between diet and treatment BPA significantly decreased the tumor multiplicity compared with other interaction groups. But no significantly contribution was from main effect treatment. And none of the variables could stay in the final model after conducting model selection procedure. The overall zero-inflation model was not significant dut to the giant standard errors for inflated-BPA, inflated-DEHP, and inflated-LC. We are not confident with the ZINB model because of the huge standard errors and the non-significance of global hypothesis.

After fitting four different models for tumor multiplicity, a best model was tentatively selected by comparing a bunch of criteria, such as log likelihood, AIC, BIC, Deviance/DF, Pearson Chi-Square/DF, and the estimated probabilities when count equals zero. All the criteria shown that there was a big improvement for model fitting from Poisson to ZIP, and another improvement from ZIP to NB and ZINB (please refer to table 15). Furthermore, because the NB model was nested in the ZINB model, the likelihood ratio test could be conducted for model comparison. And the result suggested that there was no significant difference with respect to the detective power between NB and ZINB. Therefore, simple model NB was preferred rather than ZINB model. Indeed, the NB distribution fitted the data best according to figure 10. However, the whole NB model was not significant when explained by treatment and diet. Therefore, the NB regression cannot explain where the variability of tumor multiplicity came from. Besides, over-dispersion can be the result of excess zeros or some other causes. The ZIP model is also appropriate for such data if the over-dispersion is due to the excess zeros. In this study, the origin of the excess zeros was still unknown. Yet the ZIP model did reflect the data and presented what was going on in each group. The changes of results due to take off the excess zeros in Poisson model were statistically reasonable. Only the interactive effect was not biologically reasonable. Still, ZIP model explained the variability of tumor multiplicity with respect to the treatment and diet. The reason that NB model cannot detect the difference might be because of the bad experiment administration. The DMBA induced mammary tumor model failed. There were a bunch of rats that had never developed a tumor which was not right. ZIP model as a mixture model can reflect the data better than NB because we believe our data is a mixed data with some hidden variables that we don't know yet. In summary, although the NB distribution fitted the

data best, the ZIP model reflected the data best and explained what was going on in this data mostly.

4.2 Compare to Previous literature

The results of the carcinogenic effect for BPA was consistent with the previous work [21, 22]. Due to the multitude of sources for potential exposure suggests BPA exposure in the general population [67-71], more attention needed to be paid on the contamination of BPA in the drinking water or other food containers and medical devices.

The results for DEHP were consistent with the previous work as well. It has been shown that DEHP causes malignant hepatocellular tumors in rats and mice of both sexes [72]. The U.S. Agency for Toxic Substances and Disease Registry has determined that DEHP "may be reasonably anticipated to be a carcinogen" based on animal data. However, it was lack of evidence for the carcinogenic effect of the combinations of BPA and DEHP. Our study report the significant results firstly.

The potential role of genistein, one of several known soy isoflavones, in breast cancer treatment or prevention has been extensively investigated following the early demonstration in classical animal models of chemically-induced breast cancer that dietary soy protein containing isoflavones was chemopreventive[61, 73]. In contrast, studies in the athymic mouse model of transplanted human MCF-7 breast cancer cells found that addition of genistein to the diet led to a rapid increase in tumor growth[74]. This issue has been controversial despite a total lack of supportive clinical evidence[40]. In our study, genistein did not show any chemopreventive effect in rats.

4.3 limitations of the study

Although this study provided important evidence about the effect of these four kinds of compounds on tumor latency, burden and multiplicity, there were still few limitations on experimental design and statistical inference methods.

The first problem that could introduce bias for our study belonged to the experimental design. The animal was not died for cancer naturally, but was sacrificed by the conductor once the criteria for scarifying meet. In this case, there would be some bias from the conductor, and the work time since they only worked from Monday to Friday. Besides, they keep feeding those animals after the endpoint date (300 study days) to collect more information for the study. This was a violation for the principle of the experiment.

Another problem came from the DMBA induced rat mammary tumor model. All the animals were exposed by DMBA by lavage, which was supposed to induce a low number of mammary tumors. Note that, however, there were as high as one third of animal that had never developed a tumor by the endpoint date. It indicates that there might be something wrong with the DMBA induced rat mammary tumor model. Some of the animals might not be exposed to enough DMBA to develop tumor. That is why the zero-inflated model was introduced. Since the zero-inflated model did not fit this data better than negative binomial, the problem for DMBA model had still not been solved.

There is a limitation for statistical methods as well. The mammary tumor took time to be detected. Some of the animals might die before some / all induced tumors reach a detectable size. If it was true, the response in this study, latency, burden, and multiplicity could all be affected. And hence the statistical analysis might need to be adjusted to account for uncertainty in the

number of induced tumors. Some methods may correct this problem by using specific method [75-77]. Due to very limit research time, however, this improvement can be done in the future. Also the improvement for zero-inflated negative binomial model can be developed if time permit.

Chapter 5: Conclusion

5.1 Implications of results

For answer the first questions at the beginning of the thesis, all the chemical compounds, including BPA, DEHP, Low combination and High combination, decreased the time for developing first tumor but did not have any significant effect on the survival time. We will not jump to a conclusion that the DEHP could decrease the tumor multiplicity from the results of ZIP model because first this data was not reliable statistically and second it was not reasonable biologically. The combinations of BPA and DEHP decreased tumor latency mostly, no matter low dosage or high dosage, followed by the single chemical effect. Moreover we are exposed to the chemical compounds, BPA, DEHP and their combination in our daily life widely. So we need to pay more attention to the contaminated water, plastic food container or medical devises, and the most important thing is to keep the infants away from these potential toxicities.

On the other hand, genistein and all the interactions did not show any chemopreventive effect on tumor latency, burden and multiplicity in this study. Therefore, by drinking soymilk or taking genistein complement may not help to prevent the breast cancer.

5.2 Future research

According to the limitation and the unsolved problems of this study, a series of research may need to be explored in the future. In order to explore the solve the zero inflated issue, DMBA induced model may need be redone from the very beginning and make sure that there is

no other hidden variables affect model. If the rat model cannot be redone, factor analysis or clustering analysis may need to be included to reveal the hidden variable. Also, some methods that can correct the uncertainty of the tumor occurrence can be developed in the future study.

Reference

- [1] G. Casella and R. L. Berger, *Statistical inference*, 2001.
- [2] H. Joe and R. Zhu, "Generalized Poisson Distribution: the Property of Mixture of Poisson and Comparison with Negative Binomial Distribution," *Biometrical Journal*, vol. 47, pp. 219-229, 2005.
- [3] J. A. Nelder and R. W. Wedderburn, "Generalized linear models," *Journal of the Royal Statistical Society. Series A (General)*, pp. 370-384, 1972.
- [4] P. McCullagh and J. A. Nelder, *Generalized linear models* vol. 37: Chapman & Hall/CRC, 1989.
- [5] A. Agresti, *An introduction to categorical data analysis*, Second ed. vol. 423: Wiley-Interscience, 2007.
- [6] P. D. Allison, *Logistic regression using the SAS system: theory and application*: SAS Publishing, 1999.
- [7] A. C. Cameron and P. K. Trivedi, *Regression analysis of count data* vol. 30: Cambridge University Press, 1998.
- [8] M. A. Newton and D. I. Hastie, "Assessing Poisson variation of intestinal tumour multiplicity in mice carrying a Robertsonian translocation," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 55, pp. 123-138, 2005.
- [9] W. H. Greene, "Accounting for excess zeros and sample selection in Poisson and negative binomial regression models," 1994.
- [10] N. Mantel, "Evaluation of survival data and two new rank order statistics arising in its consideration," *Cancer Chemotherapy Reports. Part 1*, vol. 50, p. 163, 1966.
- [11] R. Peto and J. Peto, "Asymptotically efficient rank invariant test procedures," *Journal of the Royal Statistical Society. Series A (General)*, pp. 185-207, 1972.
- [12] D. Collett, *Modelling survival data in medical research*. London: Chapman & Hall, 1994.
- [13] D. R. Cox, "Regression models and life tables (with discussion)," *Journal of the Royal Statistical Society*, vol. 34, pp. 187-220, 1972.
- [14] I. R. Hart, "Perspective: tumour spread—the problems of latency," *The Journal of Pathology*, vol. 187, pp. 91-94, 1999.
- [15] J. A. Aguirre-Ghiso, "Models, mechanisms and clinical evidence for cancer dormancy," *Nature Reviews Cancer*, vol. 7, pp. 834-846, 2007.
- [16] T. G. Karrison, D. J. Ferguson, and P. Meier, "Dormancy of mammary carcinoma after mastectomy," *Journal of the National Cancer Institute*, vol. 91, pp. 80-85, 1999.
- [17] J. Pfitzenmaier, W. J. Ellis, E. W. Arfman, S. Hawley, P. O. Mclaughlin, P. H. Lange, and R. L. Vessella, "Telomerase activity in disseminated prostate cancer cells," *BJU International*, vol. 97, pp. 1309-1313, 2006.
- [18] D. Weckermann, P. Muller, F. Wawroschek, R. Harzmann, G. Riethmuller, and G. Schlimok, "Disseminated cytokeratin positive tumor cells in the bone marrow of patients with prostate cancer: detection and prognostic value," *The Journal of Urology*, vol. 166, pp. 699-704, 2001.
- [19] A. S. Tsao, E. S. Kim, and W. K. Hong, "Chemoprevention of cancer," *CA: A Cancer Journal for Clinicians*, vol. 54, pp. 150-180, 2009.
- [20] B. Mafuvadze, I. Benakanakere, F. R. L. Pérez, C. Besch-Williford, M. R. Ellersieck, and S. M. Hyder, "Apigenin prevents development of medroxyprogesterone acetate-

- accelerated 7, 12-dimethylbenz (a) anthracene-induced mammary tumors in Sprague–Dawley rats," *Cancer Prevention Research*, vol. 4, pp. 1316-1324, 2011.
- [21] B. Singh, S. M. Mense, N. K. Bhat, S. Putty, W. A. Guthiel, F. Remotti, and H. K. Bhat, "Dietary quercetin exacerbates the development of estrogen-induced breast tumors in female ACI rats," *Toxicology and Applied Pharmacology*, vol. 247, pp. 83-90, 2010.
- [22] A. M. Betancourt, I. A. Eltoum, R. A. Desmond, J. Russo, and C. A. Lamartiniere, "In utero exposure to bisphenol A shifts the window of susceptibility for mammary carcinogenesis in the rat," *Environmental Health Perspectives*, vol. 118, p. 1614, 2010.
- [23] A. Y. Yakovlev, A. D. Tsodikov, and B. Asselain, *Stochastic models of tumor latency and their biostatistical applications* vol. 1: World Scientific Publishing Company Incorporated, 1996.
- [24] R. Wiehle, D. Lantvit, T. Yamada, and K. Christov, "CDB-4124, a progesterone receptor modulator, inhibits mammary carcinogenesis by suppressing cell proliferation and inducing apoptosis," *Cancer Prevention Research*, vol. 4, pp. 414-424, 2011.
- [25] S. Jagannath, W. S. Velasquez, S. L. Tucker, L. M. Fuller, P. W. McLaughlin, J. T. Manning, L. B. North, and F. C. Cabanillas, "Tumor burden assessment and its implication for a prognostic model in advanced diffuse large-cell lymphoma," *Journal of Clinical Oncology*, vol. 4, pp. 859-865, 1986.
- [26] R. A. Evans, "The tumor burden of locally recurrent breast cancer is a neglected prognostic factor," *The American Journal of Surgery*, vol. 171, pp. 445-448, 1996.
- [27] W. D. Stein, H. Huang, M. Menefee, M. Edgerly, H. Kotz, A. Dwyer, J. Yang, and S. E. Bates, "Other Paradigms: Growth Rate Constants and Tumor Burden Determined Using Computed Tomography Data Correlate Strongly With the Overall Survival of Patients With Renal Cell Carcinoma," *Cancer Journal*, vol. 15, pp. 441-447, Sep-Oct 2009.
- [28] E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *Journal of the American statistical association*, vol. 53, pp. 457-481, 1958.
- [29] E. A. Gehan, "A generalized Wilcoxon test for comparing arbitrarily singly-censored samples," *Biometrika*, vol. 52, pp. 203-223, 1965.
- [30] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 187-220, 1972.
- [31] M. P. Waalkes, J. M. Ward, and B. A. Diwan, "Induction of tumors of the liver, lung, ovary and adrenal in adult mice after brief maternal gestational exposure to inorganic arsenic: promotional effects of postnatal phorbol ester exposure on hepatic and pulmonary, but not dermal cancers," *Carcinogenesis*, vol. 25, pp. 133-141, 2004.
- [32] P. Emmelot and E. Scherer, "Multi-hit kinetics of tumor formation, with special reference to experimental liver and human lung carcinogenesis and some general conclusions," *Cancer Research*, vol. 37, pp. 1702-1708, 1977.
- [33] G. Casella and R. L. Berger, "Statistical inference," p. 92, 2001.
- [34] A. R. Moser, W. F. Dove, K. A. Roth, and J. I. Gordon, "The Min (multiple intestinal neoplasia) mutation: its effect on gut epithelial cell differentiation and interaction with a modifier system," *The Journal of Cell Biology*, vol. 116, pp. 1517-1526, 1992.
- [35] A. Moser, H. C. Pitot, and W. F. Dove, "A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse," *Science (New York, NY)*, vol. 247, p. 322, 1990.
- [36] K. A. Gould, C. Luongo, A. R. Moser, M. K. McNeley, N. Borenstein, A. Shedlovsky, W. F. Dove, K. Hong, W. F. Dietrich, and E. S. Lander, "Genetic evaluation of candidate

- genes for the Mom1 modifier of intestinal neoplasia in mice," *Genetics*, vol. 144, pp. 1777-1785, 1996.
- [37] H. Nagase, J. H. Mao, and A. Balmain, "A subset of skin tumor modifier loci determines survival time of tumor-bearing mice," *Proceedings of the National Academy of Sciences*, vol. 96, pp. 15032-15037, 1999.
- [38] N. R. Drinkwater and J. H. Klotz, "Statistical methods for the analysis of tumor multiplicity data," *Cancer Research*, vol. 41, pp. 113-119, 1981.
- [39] D. B. Dunson and J. K. Haseman, "Modeling Tumor Onset and Multiplicity Using Transition Models with Latent Variables," *Biometrics*, vol. 55, pp. 965-970, 1999.
- [40] N. M. Brown, C. A. Belles, S. L. Lindley, L. Zimmer-Nechemias, D. P. Witte, M.-O. Kim, and K. D. R. Setchell, "Mammary gland differentiation by early life exposure to enantiomers of the soy isoflavone metabolite equol," *Food and Chemical Toxicology*, vol. 48, pp. 3042-3050, 2010.
- [41] S. Jenkins, J. Wang, I. Eltoum, R. Desmond, and C. A. Lamartiniere, "Chronic oral exposure to bisphenol A results in a nonmonotonic dose response in mammary carcinogenesis and metastasis in MMTV-erbB2 mice," *Environmental Health Perspectives*, vol. 119, p. 1604, 2011.
- [42] H. Witschi, I. Espiritu, S. T. Dance, and M. S. Miller, "A mouse lung tumor model of tobacco smoke carcinogenesis," *Toxicological Sciences*, vol. 68, pp. 322-330, Aug 2002.
- [43] R. P. Singh, G. Deep, M. Chittechath, M. Kaur, L. D. Dwyer-Nield, A. M. Malkinson, and R. Agarwal, "Effect of silibinin on the growth and progression of primary lung tumors in mice," *Journal of the National Cancer Institute*, vol. 98, pp. 846-855, Jun 2006.
- [44] C. C. Conaway, C. X. Wang, B. Pittman, Y. M. Yang, J. E. Schwartz, D. F. Tian, E. J. McIntee, S. S. Hecht, and F. L. Chung, "Phenethyl isothiocyanate and sulforaphane and their N-acetylcysteine conjugates inhibit malignant progression of lung adenomas induced by tobacco carcinogens in A/J mice," *Cancer research*, vol. 65, pp. 8548-8557, Sep 2005.
- [45] L. Garcia-Segura, F. Dioloz-Bojda, V. Lenoir, F. Naftolin, and B. Kerdelhu é "Estrogen-like effects of the mammary carcinogen 7, 12-dimethylbenz (< i> a</i>) anthracene on hypothalamic neuronal membranes," *Brain research bulletin*, vol. 28, pp. 625-628, 1992.
- [46] C. Pasqualini, A. Sarrieau, M. Dussaillant, M. Corbani, F. Bojda-Diolez, W. Rost ène, and B. Kerdelhu é "Estrogen-like effects of 7, 12-dimethylbenz (a) anthracene on the female rat hypothalamo-pituitary axis," *Journal of steroid biochemistry*, vol. 36, pp. 485-491, 1990.
- [47] S. Jenkins, A. M. Betancourt, J. Wang, and C. A. Lamartiniere, "Endocrine-active chemicals in mammary cancer causation and prevention," *Journal of Steroid Biochemistry & Molecular Biology*, vol. 129, pp. 191-200, 2012.
- [48] Y. B. Wetherill, B. T. Akingbemi, J. Kanno, J. A. McLachlan, A. Nadal, C. Sonnenschein, C. S. Watson, R. T. Zoeller, and S. M. Belcher, " In vitro molecular mechanisms of bisphenol A action," *Reproductive Toxicology*, vol. 24, pp. 178-198, 2007.
- [49] J. Autian, "Toxicity and health threats of phthalate esters: review of the literature," *Environmental Health Perspectives*, vol. 4, p. 3, 1973.
- [50] B. Davis, R. Maronpot, and J. Heindel, "Di-(2-ethylhexyl) phthalate suppresses estradiol and ovulation in cycling rats," *Toxicology and applied pharmacology*, vol. 128, pp. 216-223, 1994.

- [51] J. M. Peters, M. W. Taubeneck, C. L. Keen, and F. J. Gonzalez, "Di (2-ethylhexyl) phthalate induces a functional zinc deficiency during pregnancy and teratogenesis that is independent of peroxisome proliferator-activated receptor- α ," *Teratology*, vol. 56, pp. 311-316, 1997.
- [52] J. M. Ward, J. M. Peters, C. M. Perella, and F. J. Gonzalez, "Receptor and nonreceptor-mediated organ-specific toxicity of di (2-ethylhexyl) phthalate (DEHP) in peroxisome proliferator-activated receptor α -null mice," *Toxicologic Pathology*, vol. 26, pp. 240-246, 1998.
- [53] X. O. Shu, F. Jin, Q. Dai, W. Wen, J. D. Potter, L. H. Kushi, Z. Ruan, Y. T. Gao, and W. Zheng, "Soyfood intake during adolescence and subsequent risk of breast cancer among Chinese women," *Cancer Epidemiology Biomarkers & Prevention*, vol. 10, pp. 483-488, 2001.
- [54] A. H. Wu, R. G. Ziegler, A. Nomura, D. W. West, L. N. Kolonel, P. L. Horn-Ross, R. N. Hoover, and M. C. Pike, "Soy intake and risk of breast cancer in Asians and Asian Americans," *The American Journal of Clinical Nutrition*, vol. 68, pp. 1437S-1443S, 1998.
- [55] P. M. Martin, K. B. HORWITZ, D. S. RYAN, and W. L. McGUIRE, "Phytoestrogen interaction with estrogen receptors in human breast cancer cells," *Endocrinology*, vol. 103, pp. 1860-1867, 1978.
- [56] D. T. Zava and G. Duwe, "Estrogenic and antiproliferative properties of genistein and other flavonoids in human breast cancer cells in vitro," 1997.
- [57] H. Wei, R. Bowen, Q. Cai, S. Barnes, and Y. Wang, "Antioxidant and antipromotional effects of the soybean isoflavone genistein," in *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, NY)*, 1995, pp. 124-130.
- [58] T. Akiyama, J. Ishida, S. Nakagawa, H. Ogawara, S. Watanabe, N. Itoh, M. Shibuya, and Y. Fukami, "Genistein, a specific inhibitor of tyrosine-specific protein kinases," *Journal of Biological Chemistry*, vol. 262, pp. 5592-5595, 1987.
- [59] C. Wang and M. S. Kurzer, "Effects of phytoestrogens on DNA synthesis in MCF - 7 cells in the presence of estradiol or growth factors," 1998.
- [60] T. Fotsis, M. Pepper, H. Adlercreutz, G. Fleischmann, T. Hase, R. Montesano, and L. Schweigerer, "Genistein, a dietary-derived inhibitor of in vitro angiogenesis," *Proceedings of the National Academy of Sciences*, vol. 90, pp. 2690-2694, 1993.
- [61] C. A. Lamartiniere, J. B. Moore, N. M. Brown, R. Thompson, M. J. Hardin, and S. Barnes, "Genistein suppresses mammary cancer in rats," *Carcinogenesis*, vol. 16, pp. 2833-2840, 1995.
- [62] W. B. Murrill, N. Brown, J. Zhang, P. Manzollillo, S. Barnes, and C. Lamartiniere, "Prepubertal genistein exposure suppresses mammary cancer and enhances gland differentiation in rats," *Carcinogenesis*, vol. 17, p. 1451, 1996.
- [63] L. Hilakivi-Clarke, I. Onojafe, M. Raygada, E. Cho, T. Skaar, I. Russo, and R. Clarke, "Prepubertal exposure to zearalenone or genistein reduces mammary tumorigenesis," *British Journal of Cancer*, vol. 80, p. 1682, 1999.
- [64] J. K. Day, C. Besch-Williford, T. R. McManm, M. G. Hufford, D. B. Lubahn, and R. S. MacDonald, "Dietary genistein increased DMBA-induced mammary adenocarcinoma in wild-type, but not ER α KO, mice," *Nutrition and Cancer*, vol. 39, pp. 226-232, 2001.

- [65] P. Kijkuokool, I. S. Parhar, and S. Malaivijitnond, "Genistein enhances N-nitrosomethylurea-induced rat mammary tumorigenesis," *Cancer Letters*, vol. 242, pp. 53-59, 2006.
- [66] N. M. Brown, P. A. Manzollilo, J. Zhang, J. Wang, and C. A. Lamartiniere, "Prenatal TCDD and predisposition to mammary cancer in the rat," *Carcinogenesis*, vol. 19, pp. 1623-1629, 1998.
- [67] A. M. Calafat, Z. Kuklennyik, J. A. Reidy, S. P. Caudill, J. Ekong, and L. L. Needham, "Urinary Concentrations of Bisphenol A and 4-Nonylphenol in a Human Reference Population," *Environmental Health Perspectives*, vol. 113, pp. 391-395, 2005.
- [68] A. M. Calafat, J. Weuve, X. Ye, L. T. Jia, H. Hu, S. Ringer, K. Huttner, and R. Hauser, "Exposure to bisphenol A and other phenols in neonatal intensive care unit premature infants," *Environmental Health Perspectives*, vol. 117, p. 639, 2009.
- [69] A. M. Calafat, X. Ye, L. Y. Wong, J. A. Reidy, and L. L. Needham, "Exposure of the US population to bisphenol A and 4-tertiary-octylphenol: 2003–2004," *Environmental Health Perspectives*, vol. 116, p. 39, 2008.
- [70] J. L. Carwile, H. T. Luu, L. S. Bassett, D. A. Driscoll, C. Yuan, J. Y. Chang, X. Ye, A. M. Calafat, and K. B. Michels, "Polycarbonate bottle use and urinary bisphenol A concentrations," *Environmental Health Perspectives*, vol. 117, p. 1368, 2009.
- [71] M. S. Wolff, S. L. Teitelbaum, G. Windham, S. M. Pinney, J. A. Britton, C. Chelimo, J. Godbold, F. Biro, L. H. Kushi, and C. M. Pfeiffer, "Pilot study of urinary biomarkers of phytoestrogens, phthalates, and phenols in girls," *Environmental Health Perspectives*, vol. 115, p. 116, 2007.
- [72] W. M. Kluwe, "Overview of phthalate ester pharmacokinetics in mammalian species," *Environmental Health Perspectives*, vol. 45, p. 3, 1982.
- [73] C. A. Lamartiniere, M. S. Cotroneo, W. A. Fritz, J. Wang, R. Mentor-Marcel, and A. Elgavish, "Genistein chemoprevention: timing and mechanisms of action in murine mammary and prostate," *The Journal of nutrition*, vol. 132, pp. 552S-558S, 2002.
- [74] C.-Y. Hsieh, R. C. Santell, S. Z. Haslam, and W. G. Helferich, "Estrogenic effects of genistein on the growth of estrogen receptor-positive human breast cancer (MCF-7) cells in vitro and in vivo," *Cancer research*, vol. 58, pp. 3833-3838, 1998.
- [75] D. B. Dunson and G. E. Dinse, "Distinguishing Effects on Tumor Multiplicity and Growth Rate in Chemoprevention Experiments," *Biometrics*, vol. 56, pp. 1068-1075, 2000.
- [76] L. s. Freedman, D. N. Midthune, C. C. Brown, V. Steele, and G. J. Kelloff, "Statistical analysis of animal cancer chemoprevention experiments," *Biometrics*, pp. 259-268, 1993.
- [77] S. M. Kokoska, "The analysis of cancer chemoprevention experiments," *Biometrics*, pp. 525-534, 1987.

Appendix 1: Tables and Figures

Table 16: Model Selection of Poisson Model for Tumor Multiplicity

| Model | Variables | DF | Log Likelihood | AIC | BIC | Comparison | Likelihood ratio test | P=.05 Critical Value |
|-------|---------------------------|----|----------------|---------|--------|------------|-----------------------|----------------------|
| 0 | Intercept only | 0 | -1450.81 | 2903.61 | 2907.4 | -- | -- | -- |
| 1 | Treatment | 4 | -1422.07 | 2854.15 | 2873.1 | 0 vs 1 | 57.48*** | 9.49 |
| 2 | Treatment, Treatment*Diet | 9 | -1401.59 | 2823.17 | 2861.1 | 1 vs 2 | 40.96*** | 11.07 |

Note: *** P < .0001

Table 17: Model Selection of NB Model for Tumor Multiplicity

| Model | Variables | DF | Log Likelihood | AIC | BIC | Comparison | Likelihood ratio test | P=.05 Critical Value |
|-------|---------------------------|----|----------------|---------|--------|------------|-----------------------|----------------------|
| 0 | Intercept only | 1 | -940.866 | 1885.73 | 1893.3 | -- | -- | -- |
| 1 | Treatment | 5 | -936.582 | 1885.16 | 1907.9 | 0 vs 1 | 8.568 | 9.49 |
| 2 | Treatment, Treatment*Diet | 10 | -932.986 | 1887.97 | 1929.7 | 1 vs 2 | 7.192 | 11.07 |

Table 18: Variable Selection for Poisson Given in ZIP Model for Tumor Multiplicity

| Model | Variables | DF | Log Likelihood | AIC | BIC | Comparison | Likelihood ratio test | P=.05 Critical Value |
|-------|---------------------------|----|----------------|------|------|------------|-----------------------|----------------------|
| 0 | Intercept only | 4 | -1062.514 | 2137 | 2160 | -- | -- | -- |
| 1 | Treatment | 8 | -1049.35 | 2119 | 2157 | 0 vs 1 | 26.328*** | 9.49 |
| 2 | Treatment, Treatment*Diet | 13 | -1037.16 | 2104 | 2161 | 1 vs 2 | 24.38*** | 11.07 |

Note: *** P < .0001

Table 19: Variable Selection for NB Given in ZINB Model for Tumor Multiplicity

| Model | Variables | DF | Log Likelihood | AIC | BIC | Comparison | Likelihood ratio test | P=.05 Critical Value |
|-------|---------------------------|----|----------------|------|------|------------|-----------------------|----------------------|
| 0 | Intercept only | 5 | -936.233 | 1886 | 1913 | -- | -- | -- |
| 1 | Treatment | 9 | -932.963 | 1888 | 1930 | 0 vs 1 | 6.540 | 9.49 |
| 2 | Treatment, Treatment*Diet | 14 | -929.234 | 1890 | 1951 | 1 vs 2 | 7.457 | 11.07 |

Figure 11: Boxplot for Tumor Latency by Treatment, Adjusting Diet

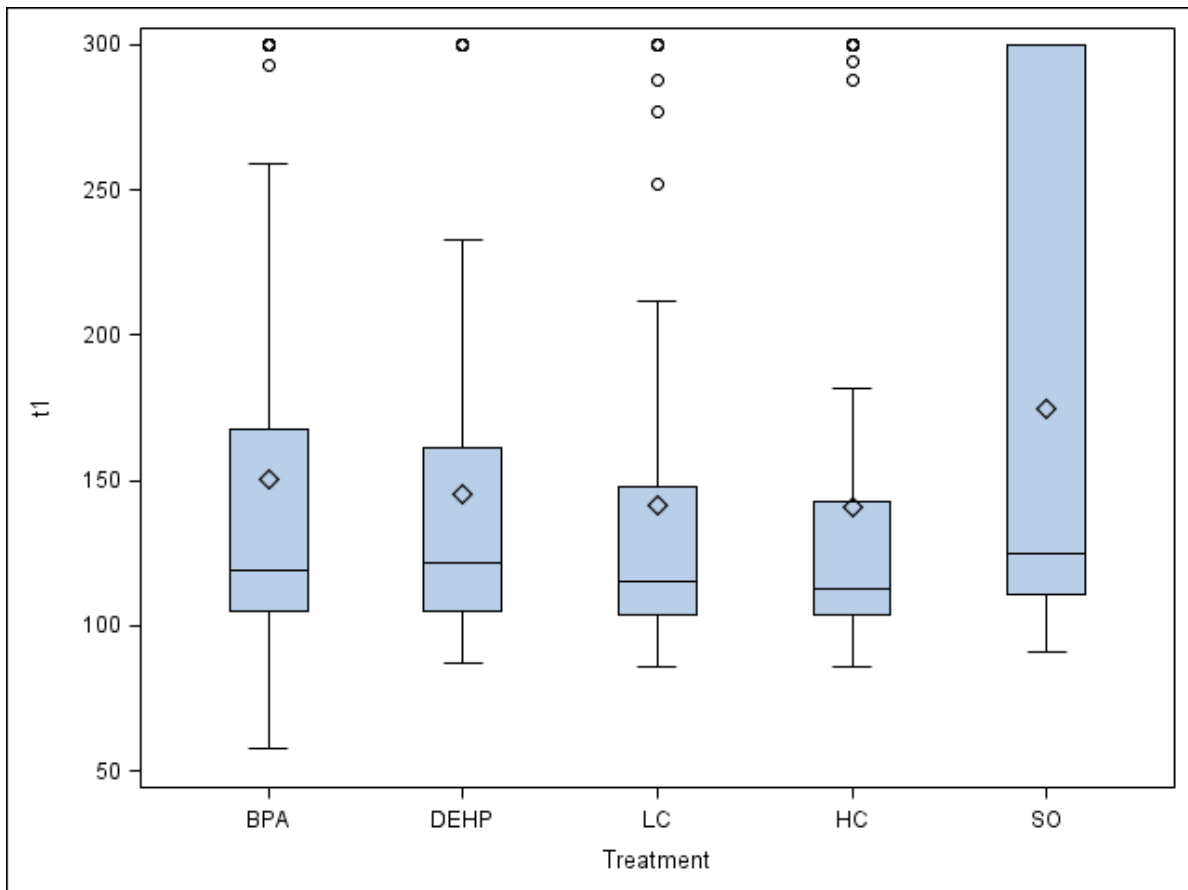


Figure 12: Boxplot for Tumor Burden by Treatment, Adjusting Diet

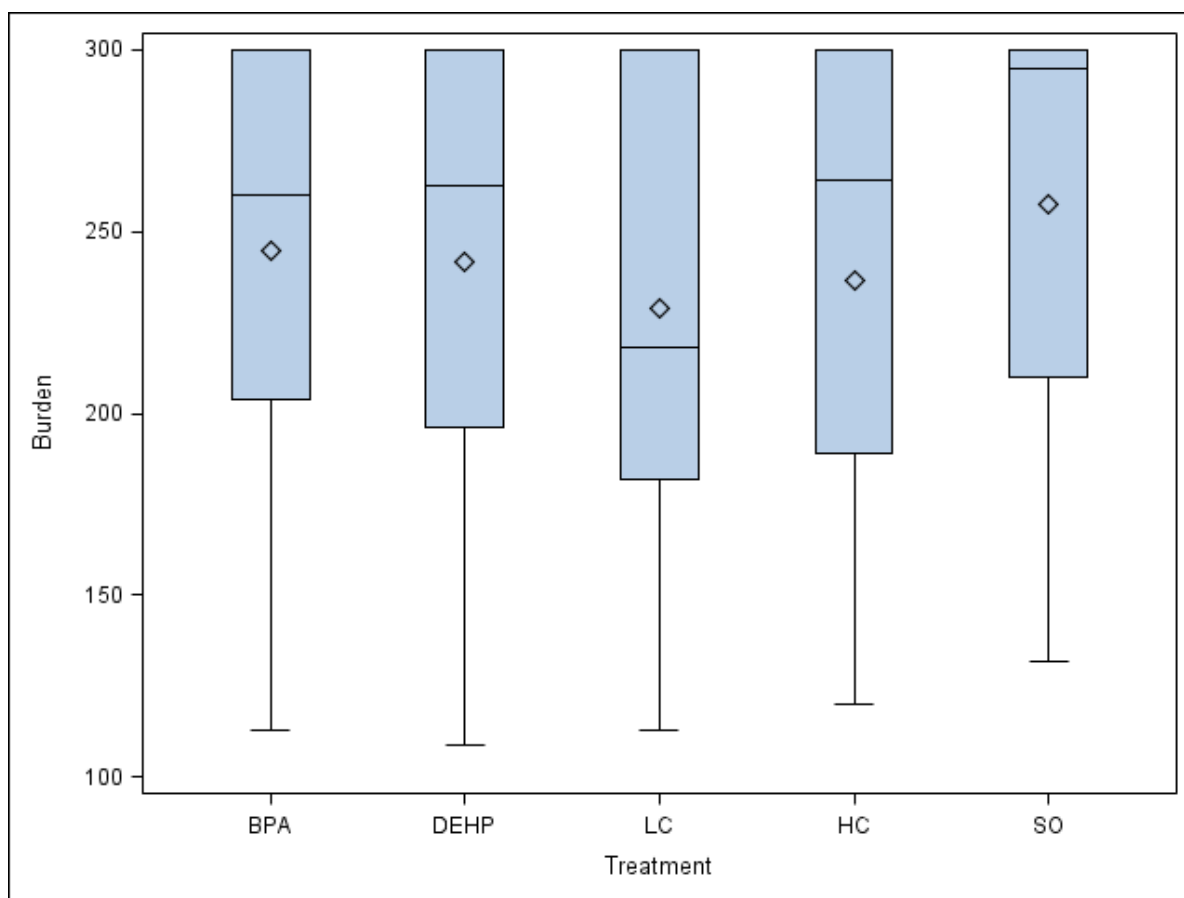


Figure 13: Boxplot for Tumor Multiplicity by Treatment, Adjusting Diet

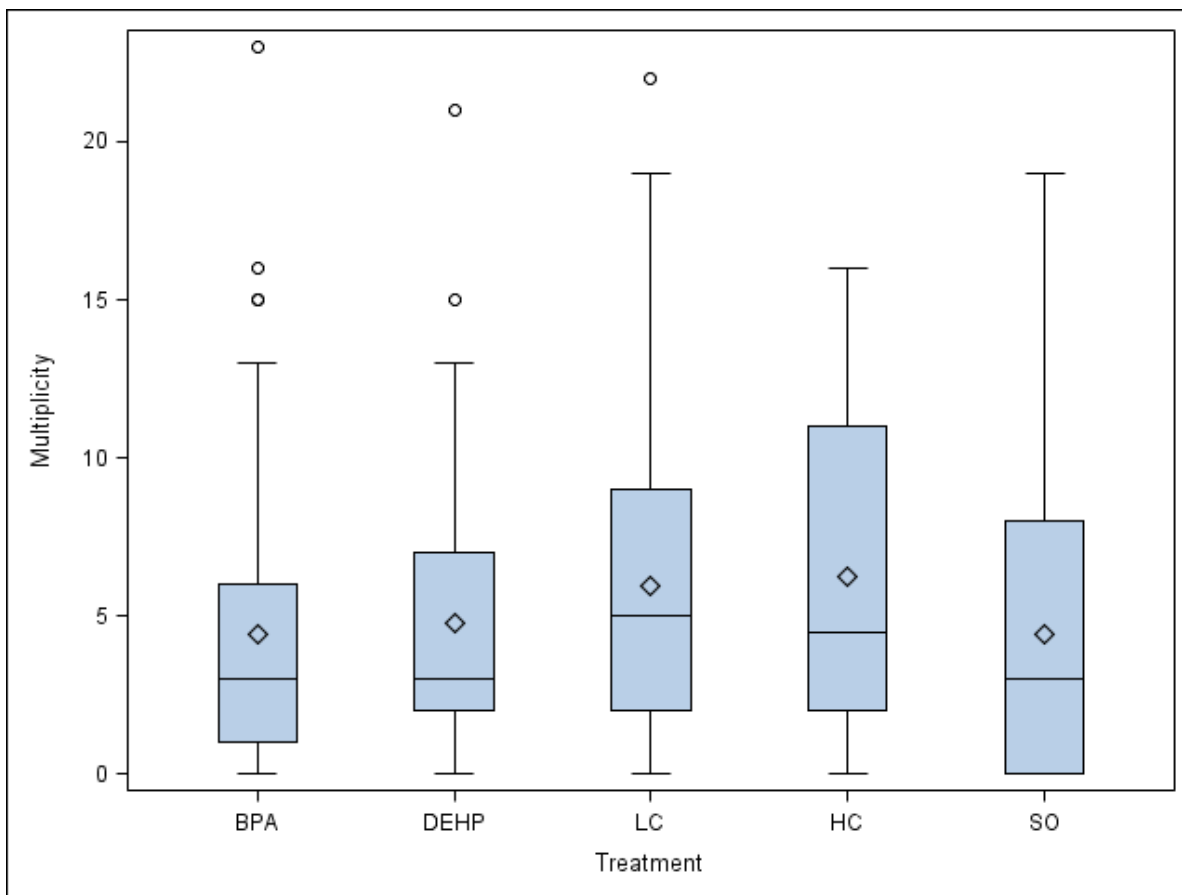
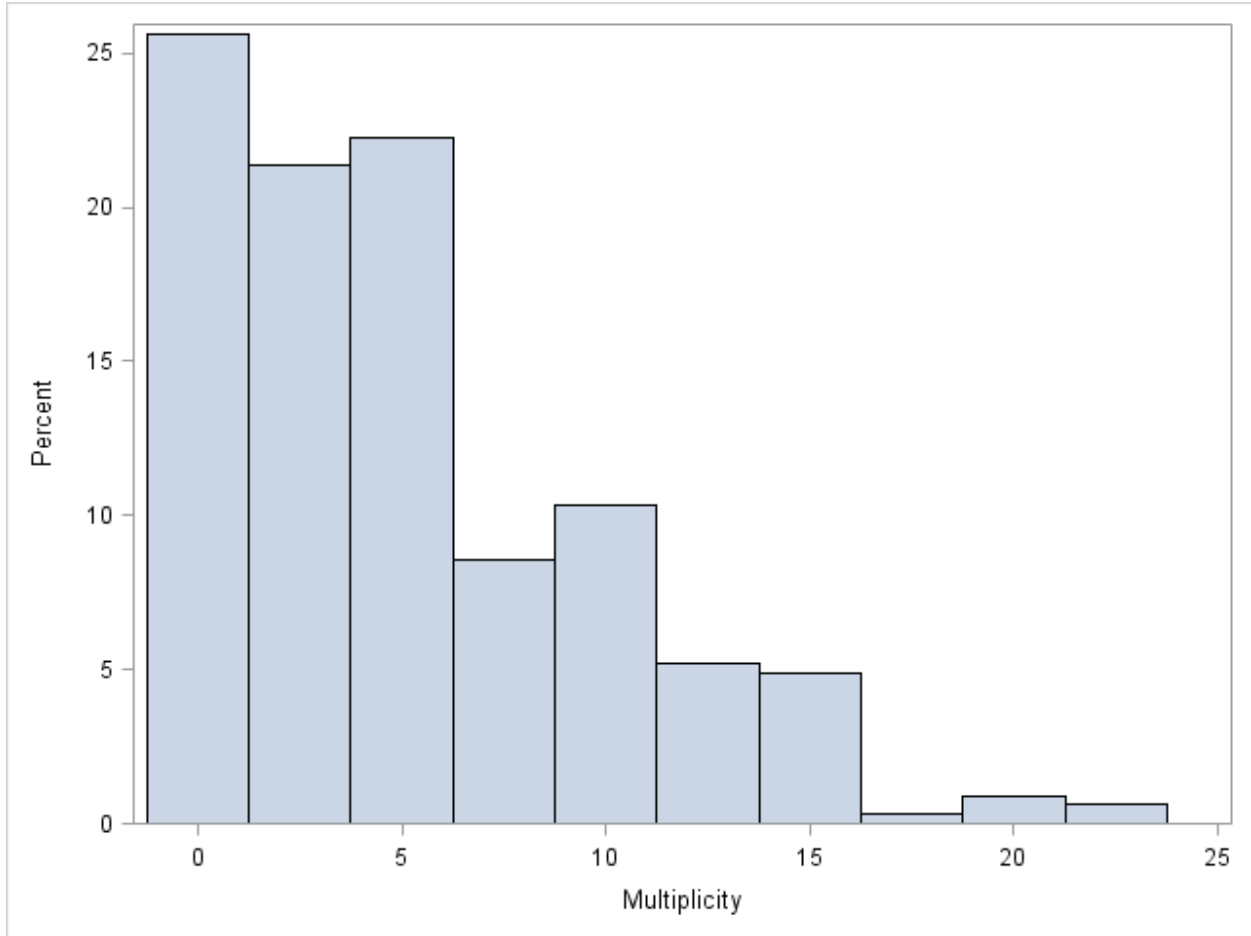


Figure 14: Histogram of Overall Tumor Multiplicity



Appendix 2: Partial SAS Code

```
proc import out=data0
  datafile='D:\*****'
  dbms=csv replace;
  getnames=yes;
run;

proc format;
  value treat 1='BPA' 2='DEHP' 3='LC' 4='HC' 5='SO';run;
proc format;
  value diet 1='Gen' 0='Con';run;
proc format;
  value tgp 1='BPA' 2='DEHP' 3='LC' 4='HC' 5='SO' 6='SO+GEN' 7='BPA+GEN'
8='DEHP+GEN' 9='LC+GEN' 10='HC+GEN';run;
```



```

data data0; set data0;
    Treat= 1*(group in (1,7)) + 2*(group in (2,8)) + 3*(group in (3,9))+
4*(group in (4,10)) + 5*(group in (5,6));
    Diet=(group > 5);
    format treat treat. group tgp. diet diet.;
    drop treatment; rename treat=Treatment;
run;

data data300; set datacheck; cutoff_day=300; if days_on_study >=0;
array palpable_t{23} palpable_t1-palpable_t23;
array tcensor{23} tcensor1-tcensor23;
array t{23} t1-t23;
do i=1 to 23;
tcensor{i}=(palpable_t{i}-dob >cutoff_day or palpable_t{i}=.);
t{i}=min(palpable_t{i}-dob, cutoff_day);
end;
Burden=min(cutoff_day,rdays_on_study);
T2tumor= t2-t1;
t2censor=tcensor2-tcensor1;
Lnb=log(Burden);
LNT1=LOG(T1);
logbw=log(body_weight);
loguw=log(uterine_weight);
censor300=(rdays_on_study > cutoff_day);
Weight_ratio=body_weight/uterine_weight;
Multiplicity=N(OF t1-t23)-SUM(OF tcensor1-tcensor23);
Tumor_rate=Multiplicity/rdays_on_study*300;
BPA=(treatment=1);
DEHP=(treatment=2);
LC=(treatment=3);
HC=(treatment=4);
BPAGEN=BPA*DIET;
DEHPGEN=DEHP*DIET;
LCGEN=LC*DIET;
HCGEN=HC*DIET;
KEEP animal_id cage dob treatment group body_weight uterine_weight rsac_date
    rdays_on_study total_tumors total_tumor_volume tumor_burden LNT1 logbw
loguw
    t1-t23 burden T2tumor t2censor lnb cutoff_day censor300 diet tcensor1-
tcensor23
    weight_ratio multiplicity tumor_rate Int BPA DEHP LC HC BPAGEN DEHPGEN
LCGEN HCGEN;
run;

PROC MEANS data=data300 maxdec=2 N median mean stderr min max noprint;
    CLASS treatment diet group;
        var burden;
        OUTPUT out=means1 n=n median=median mean=mean stderr=stderr min=min
max=max;
RUN;

proc sgplot data=data300;
vbox burden/ category=group;
run;

proc freq data=data300;

```

```

table treatment*diet/ chisq;
weight Multiplicity;
run;

proc logistic data=data300;
class treatment (ref='SO') diet (ref='Con')/ param=reference;
model tcensor1 (event='1')=treatment diet;
run;

PROC LIFETEST data=data300 plots=(s, ls, lls);
STRATA treatment; TIME t1*tcensor1(1); test diet;
title 'Lifetest for Tumor Latency in 300 Days';run;
PROC LIFETEST data=data300 plots=(s, ls, lls, h);
strata treatment; time burden*censor300(1); test diet BPAGEN DEHPGEN ;
title 'Lifetest for Tumor Burden in 300 Days';run;

PROC LIFEREG data=data300 outest=elt; CLASS treatment; MODEL
t1*tcensor1(1)=treatment diet treatment*diet;
title 'Lifereg for Tumor Latency in 300 Days';
RUN;
PROC LIFEREG data=data300; CLASS treatment; MODEL
burden*censor300(1)=treatment diet treatment*diet;
title 'Lifereg for Tumor Burden in 300 Days';
RUN;

proc phreg data=data300;class treatment;
model burden*censor300(1)= treatment diet treatment*diet/selection=forward;
run;

proc phreg data=data300;class treatment;
model burden*censor300(1)= treatment x1 X2 X3 X4;
OUTPUT OUT=Outp xbeta=xb resmart=mart resdev=dev;
x1=lnb*(treatment=1);
x2=lnb*(treatment=2);
x3=lnb*(treatment=3);
x4=lnb*(treatment=4);
proportionality_test: test x1, X2, X3, X4;
RUN;

proc gplot data=Outp;
plot (mart dev)*xb / vref=0 cframe=ligr;
symbol1 value=circle c=blue;
run;

PROC GENMOD data=data300 plots=all; class treatment (ref='SO');
model multiplicity=treatment diet treatment*diet/offset=lnb dist=NB;
title 'Multiplicity Estimated by Treatment and Diet for Negative
Binominal';
run;

proc countreg data=data00;
model multiplicity=BPA DEHP LC HC DIET BPAGEN DEHPGEN LCGEN HCGEN/offset =
lnb dist=zip;
zeromodel multiplicity~BPA DEHP LC HC DIET BPAGEN DEHPGEN LCGEN HCGEN;
title 'Countreg for ZIP model';
run;

```