2000

A Bayesian approach for assessment of therapeutic equivalence of three proportions.

Delicia Evet Carey
University of Alabama at Birmingham

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

A BAYESIAN APPROACH FOR ASSESSMENT OF THERAPEUTIC EQUIVALENCE OF THREE PROPORTIONS

by

DELICIA EVET CAREY

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2000

ABSTRACT OF DISSERTATION
GRAUATE SCHOOL, UNIVERSITY OF ALABAMA AT BIRMINGHAM

Degree Ph.D. Program Biostatistics

Name of Candidate Delicia Evet Carey

Committee Chairs Alfred A. Bartolucci and Karan P. Singh

Title A Bayesian Approach for Assessment of Therapeutic Equivalence of Three Proportions

Nonrejection of the null hypothesis when comparing two response rates in a clinical trial does not necessarily imply that the two treatments are equivalent with respect to their therapeutic effectiveness. Recently, researchers have invested much time and effort in describing situations in which therapeutic equivalence (TE) may be achieved. These have involved direct hypothesis-testing procedures and confidence interval techniques. The latter involves determining whether such an interval lies within predefined equivalent regions. The committee chairs (Al Bartolucci and Karan Singh) have published previously on this subject when the parameters of interest were from growth curve or survival distributions. In this research endeavor, the focus is to determine whether three treatments are equivalent with respect to their therapeutic effectiveness. The prior information involves the natural conjugate beta family of distributions. The primary parameter of interest is the ratio of the two binomial parameters. Limiting values of the hyperparameters of the conjugate family are used to demonstrate the robustness of the outcomes. Several equivalence regions are used to test whether equivalence has been achieved and under what conditions the attainment of equivalence may not be established. The procedure has wide applications as well in the quality control setting.

DEDICATION

This document is dedicated to the memory of my maternal grandfather affectionately known as "Daddy"; my maternal grandmother, Dear; Dotha, Dai, Eloise, Harvey, Neal, Mr. Wells, and Alfrado.

ACKNOWLEDGEMENTS

First and foremost, I give thanks, honor, and praise to my Lord and Savior, Jesus Christ. I would like to thank my advisor, Dr. Bartolucci, for his assistance and support during my graduate studies at UAB. The encouragement and support of Dr. Singh, my co-chair, is greatly appreciated. Thanks go to the other members of my committee, Dr. Roy, Dr. Weiss, and Dr. Williams. I am very appreciative of their service on my committee. Thanks to Dr. Katholi for aiding me in some very tough numerical calculations. Finally, I would like to thank my family for their undying love and support throughout my educational career.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

PREFACE

This dissertation addresses some of the fundamental issues in clinical trials research regarding equivalence. Unresolved issues pertaining to the determinance of equivalence are numerous. Some of those issues are design, analysis, and inference. However, the focus of this dissertation is on the analysis aspect of the assessment of equivalence. The proposed approach gives researchers a method to compare three proportions simultaneously. My research contribution is twofold. First is the derivation of the posterior density of the ratio of the proportions using the beta-binomial distribution. Second is the derivation of the weighted likelihood ratio.

CHAPTER 1

INTRODUCTION

Prologue

Demonstrating therapeutic equivalence of two treatments is the goal of many clinical trials. Generally speaking, the topic of this dissertation is equivalence. Equivalence studies are often needed to develop better tolerated therapies after effective treatments have been identified. Clinical trials conducted with the goal of demonstrating therapeutic equivalence have been the concern of the government, the pharmaceutical industry, academia, and the medical industry. Moreover, the concept of equivalence may include the manufacturing industry. In this arena, manufacturers are interested in providing or assuring that machines, services, products, and processes are equivalent in performance with respect to prototypes.

Many issues are still unclear and unresolved in the equivalence problem area. There are design, inference, and analysis issues to name just a few.

Previous research has considered various aspects of testing individual proportions. This discourse will extend the existing Bayesian methods for assessment of therapeutic equivalence in the context of equivalence testing. Hence, the emphasis of this writing is on both the Bayesian method for assessment of therapeutic equivalence (TE) in the binomial context and extending the theory to three proportions in clinical trials by using the beta-binomial distribution and the derivation of the weighted likelihood ratio.

Statement of the Problem

Most Phase III studies are designed with several endpoints in mind. One of these endpoints is the rate or proportion.

In a clinical trial, interest sometimes focuses on the question of which therapy produces the highest proportion of successes or responses. Response can be defined according to the criteria given in a treatment protocol. Proportions are used to depict the percentage of patients in a given trial or on a particular treatment who have certain characteristics.

Although there are numerous techniques based on classical statistical approaches, it is worth mentioning that the Bayesian approach has come to the forefront in the literature. Hence, the focus of this research problem is based on the Bayesian approach.

This dissertation focuses on the analysis aspect of equivalence. It assumes that the design aspects of the study were carried out such that all parties of concern agree with the method. As a result, a formal statement of the problem is as follows: Given a data set of three treatments, with the objective of demonstrating equivalence of proportions in an active control clinical trial, how does one analyze the data set, assuming that the underlying distribution is binomial and that the prior information about the parameters of the distribution is beta?

Research Objectives

There are several important objectives of this research. One objective of this research is to extend the methodology of Pham-Gia and Turkkan (1993) to include other ways of comparing proportions. The purpose of this objective is to give researchers greater flexibility in efficiently comparing at least three proportions simultaneously. The

Pham-Gia and Turkkan (1993) methodology is also taken a step further to the joint distribution of the ratio of proportions for two proportions. In addition, this research considers the pairwise ratio of three proportions.

Another objective of this research is to present an integration technique that can be used in integrating posterior distributions on credibility limits when using this methodology. This integration technique is shown with the aid of Mathcad, a standard calculation software. The Monte Carlo method for integration is also used. An additional objective is to derive the weighted likelihood ratio of the ratio of proportions. Finally, a further objective of this research is to provide a graphical method for the comparison of three proportions.

An Overview of Other Approaches

Over the years, many researchers have worked on the concept of therapeutic equivalence. The research has included direct hypothesis-testing procedures and confidence interval techniques. The beginning work done with the problem of showing bioequivalence of drug formulations was most often treated as a two-sided problem. The idea of this problem was to show that the effect of a new formulation did not differ substanstantially in either direction from that of a standard formulation. Blackwelder (1982) focused on bioequivalence.

Later, the two-sided problem set-up was suggested inappropriate. Blackwelder and Chang (1984) present sample size graphs for proving the null hypothesis and set up appropriate parametric regions for applying their approach. Glatstein and Makuch (1984) take the design power approach, and Hauck and Anderson (1986) propose an alternative approach in which they define equivalence to mean that actual differences lie within

some specified confidence limits. Durrleman and Simon (1990) further describe sequential monitoring of equivalence studies using repeated confidence intervals. Westlake (1972) and Metzler (1974) proposed the use of confidence intervals as a tool in data interpretation and decision making.

In the recent past, a Bayesian approach has been investigated in which the probability of a relative difference of a certain size or larger is estimated. Fluehler, Grieve, Mandallaz, Mau, and Moser (1983) proposed a Bayesian approach that involves obtaining the posterior probability that the ratio of the true means of a new and a standard formulation lies within a given interval. In addition, Selvyn, Dempster, and Hall (1981) presented an alternative Bayesian approach in which the criterion for equivalence is that the difference in the response means is less than a specified percentage of the mean of the standard, based on the posterior probability.

The Bartolucci and Singh (1993) methodology followed that of the confidence region methodology, in which they defined a general class of discrepancy measures between parameters of interest and then applied the Bayesian neighborhood null hypothesis theory to derive posterior confidence regions on those measures (Bartolucci and Singh, 1993). The authors looked at this subject when the parameters of interest were from growth curve or survival distributions.

Jefferson, Bartolucci, and Singh (1997) took this idea a step further and derived the asymptotic joint distribution of the credibility limits so that probabilistic statements could be made about the credibility limits being a subset of the specified interval or any other interval.

Berger and Hsu (1996) discuss the equivalence confidence set method for equivalence. This method is recognized as one for which the derived inference, instead of

the usual significant difference, is practical equivalence. For this method, Berger and Hsu (1996) demonstrate bioequivalence by testing the following hypothesis of this active control trial:

$$H_o : \frac{\mu_T}{\mu_R} \leq \delta_L \text{ or } \frac{\mu_T}{\mu_R} \geq \delta_U \quad (1)$$

$$H_a : \delta_L < \frac{\mu_T}{\mu_R} < \delta_U \quad (2)$$

The values δ_L and δ_U are standards set by regulatory agencies that define how "close" the drugs must be to be declared bioequivalent. The values μ_T and μ_R denote the population mean area under curve (AUC) for the test drug and the population mean AUC for the reference drug, respectively.

As stated previously, the confidence interval technique involves determining whether a confidence interval lies within predefined equivalence regions. Carey, Bartolucci, and Singh (1998) focus on binomial parameters characterizing the response from clinical trials. The approach follows that of the confidence region methodology, in which a general class of discrepancy measures between parameters of interest is defined and the Bayesian neighborhood null hypothesis theory is applied to derive posterior confidence regions on these measures. The theory of this application follows that of Bartolucci and Dickey (1977) and Dickey (1979). The use of a realistic family of prior distribution, the natural conjugate beta family of distributions, follows the technique of Birch and Bartolucci (1983) and Bartolucci, Katholi, and Birch (1992). Finally, it is determined whether the derived measure falls within an equivalence region of interest as in Hauck and Anderson (1986).

Carey, Bartolucci, and Singh (1998) considered equivalence testing with two treatments, each modeled by an appropriate binomial distribution with parameters p_1 and p_2 as expressed by Hauck and Anderson (1986),

$$-\Theta < p_1 - p_2 < \Theta, \quad \Theta > 0 \tag{3}$$

Carey, Bartolucci and Singh also applied the theory of Bartolucci and Dickey (1977) and Dickey (1979), where the neighborhood null hypothesis is

$$H_o : |\eta - \eta_o| < \Theta \tag{4}$$

where, in the present, $\eta = p_1 - p_2$ and $\theta = 0.2$. Some 1-α posterior probability bounds for η are determined. In addition, it is determined whether these bounds are in $\pm\theta$ or less in width. A choice for θ may be 0.20 (Huque, Dubey, & Fredd, 1990).

A Sketch of the Methodology of the Proposed Solution

Pham-Gia and Turkkan (1993) presented the Bayesian analysis of the difference of two populations. This approach will be discussed further in chapter 2. The approach taken in this dissertation will be an extension of Pham-Gia and Turkkan's (1993) work. This methodology will be extended to include three proportions.

The overall objective of this dissertation is to provide researchers with a way to assess therapeutic equivalence of three proportions simultaneously. The action items of the objective are as follows:

1. The posterior distribution for the ratio of two proportions at a time will be. Each combination of the two proportions will be used. This will allow the three proportions to be compared by analyzing sets of two.
2. A 100 (1- α)% confidence region of the ratio of two proportions will be derived.

3. The idea behind the ratio of two proportions will be extended in pairwise fashion to three proportions.
4. The posterior distribution will be derived using all three proportions.
5. The criteria based on a $100(1-\alpha)\%$ credible region for the proportions will be used to establish a necessary condition regarding equivalence.
6. The weighted likelihood ratio will be derived.
7. Equivalence will also be determined based on a three-dimensional measuring scale coined the "equivalence cube."

Content of Other Chapters

The remainder of the document contains chapter 2 through 7. There are appendixes and a reference list to support information within those chapters.

Chapter 2 contains a detailed review of some of the references cited in the overview section of chapter 1. It also contains a review not listed in the overview section of chapter 1. The analyses and methods in this section vary from Bayesian and frequentist approaches. In some cases the random variables are discrete, and in other cases they are continuous.

In chapter 3, an overview of therapeutic equivalence is given. Definitions of therapeutic equivalence are presented from the perspective of Blackwelder (1982), the FDA (1979), and others. The advantages of using therapeutic equivalence are discussed.

Chapter 4 contains an overview of active control clinical trials. The factors that influence clinical evaluations are presented. The appeal of the active control trial and the criticisms of that type of trial are presented. Chapter 5 contains a detailed discussion of the methodology. This involves determining the prior distribution to the development of

the credibility regions of the ratios of proportions. The actual determinance method of equivalence is presented. The discussion also includes problems encountered when integrating posterior kernels in multidimensions.

Chapter 6 is an application of the methodology to a published study. The formulations involved the treatment of cancer. In an advanced non-small-cell carcinoma of the lung, trial patients were randomized to one of three treatments (a) CAMF (cyclo-phosphamide adriamycon methotrexate with folinic acid), (b) CAP (cyclophosphamide adriamycin cisplatinum) and (c) CA (cyclophosphamide adriamycin). The methodology was also applied to test data for means of comparison.

Finally, chapter 7 summarizes the work presented in the previous chapters. It also discusses possible extensions to this research. Future research is suggested with respect to analysis.

CHAPTER 2

SELECTED DETAILED REVIEW OF RELATED LITERATURE

Introduction

The different statistical approaches to the problem of equivalence are placed in different sections of this chapter. Those selected are in chronological order with respect to the date of publication. The approaches are presented from the stating of the hypotheses to the derived test statistic. The details of the derivations are omitted. However, the reference for each approach is contained in the reference section for the reader who desires further detail.

Good (1950). As part of this research, we develop methodology and notation for making coherent inferences on parameters of sampled distributions. In particular, we will investigate the weighted likelihood ratio (WLR), also known as the Bayes' factor. This section discusses the work of I. J. Good, who actually termed Bayes' factor.

Hence, the purpose of this section is to explain the process of weighing evidence as explained by Good.

Let

$$\frac{P(H|E)}{P(H)} \propto P(E|H), \tag{5}$$

where, E is fixed and H is variable. This theorem is known as the principle of inverse probability. For most applications, E is the result of an experiment and H is the hypoth-

esis. R. A. Fisher called $P(H|E)$ the *likelihood* of H given because he wanted to avoid the use of the term Bayes' theorem.

When there are two hypotheses, denoted by H and H^c, the following exists:

$$\frac{O(H|E)}{O(H)} = \frac{P(E|H)}{P(E|\overline{H})}, \tag{6}$$

where $O(H)$ is the initial odds and $O(H|E)$ is the final odds.

The odds of H given E is given by the following equation:

$$O(H|E) = \frac{P(H|E)}{1 - P(H|E)}.$$

These odds should not be confused with betting odds.

The ratio of $P(E|H)/P(E|H)$ is the ratio of the likelihoods of H and H^c with respect to E. This case of Bayes' theorem may be stated as in the following theorem: The factor in favor of a hypothesis H is equal to the ratio of the likelihoods of H and H^c.

Dr. A. M. Turing used acoustics and electrical engineering notation to describe *weight of evidence*. From acoustics, the bel is the logarithm to the base 10 of the ratio of two intensities of sound. Likewise, if f is the factor in favor of a hypothesis, then the hypothesis has gained $\log_{10} f$ bels, which is described as the weight of evidence. Hence, "Plausibility gained = weight of evidence," where the weight of evidence is calculated in terms of the ratio of the likelihoods. A noteworthy theorem is as follows: Suppose that a series of experiments are performed, with results $E_1, E_2, \ldots, E_n$, and suppose that these are independent given H and independent given H^c. Then the resulting factor is equal to the product of the individual factors, and therefore the resulting weight of evidence is equal to the sum of the individual weight of evidence. Because of the in-

dependence conditions,

$$\frac{P(E_1 E_2 \ldots E_n|H)}{P(E_1 E_2 \ldots E_n|\overline{H})} = \frac{P(E_1|H)}{P(E_1|\overline{H})} \cdots \frac{P(E_n|H)}{P(E_n|\overline{H})},$$

where the factors are multiplicative and the weights of evidence are additive. Consider having a composite hypothesis, H, where H is expressed as the disjunction of n mutually exclusive hypotheses $H_1, H_2, \ldots, H_n$. The interest is in whether a hypothesis H is true with evidence E and some evidence H' which is taken for granted. Let

$$P(H_r|H) = P_r.$$

Then we have the following theorem: The factor in favor of H in virtue of E is equal to the "weighted average" of the partial factors. For proof see Appendix A.

The definition of expected weight of evidence is as follows: Let H be the hypothesis and $E_1, E_2, \ldots, E_n$ be exclusive results of an experiment. Then,

$$\sum_r P(E_r|\overline{H}) * \frac{P(E_r|H)}{P(E_r|\overline{H})} = \sum_r P(E_r|H) = P(\mathrm{E1}\,U\,\mathrm{E2}\,U \ldots U\,\mathrm{En}|\mathrm{H}) = 1.$$

According to Good, the expected weight of evidence for *correct* hypotheses is positive and for *wrong* hypotheses is negative.

Blackwelder (1982). Blackwelder discusses the question of whether a new therapy is as effective as a standard therapy. According to the author, in determining whether an experimental therapy is as effective as a standard therapy, a test of the conventional null hypothesis that the two treatments have equal effects leads to logical difficulties that can be overcome with a different formulation of the hypothesis. The logical difficulty is that in this case the desired result of failing to reject the null hypothesis cannot be accomplished. In addition, it is inappropriate, regardless of sample

size, to base a conclusion that therapies are equivalent on whether the observed significance level for the null hypothesis of equality is larger than some arbitrary small value. In a study designed to show equivalence of therapies, the quantity, δ, is sufficiently small so that the therapies are considered equivalent for practical purpose, (i.e., if the difference $\pi_s - \pi_e$ is less than the minimum difference of practical interest). The true success probabilities, π_s and π_e, are for the standard and experimental therapy, respectively. The null hypothesis would be as follows:

$$H_o : \pi_s \geq \pi_e + \delta. \tag{7}$$

The alternative hypothesis would be as follows:

$$H_a : \pi_s \leq \pi_e + \delta. \tag{8}$$

Assuming sufficiently large samples to justify the normal approximation to the binomial, the test statistic is now

$$z = \frac{p_s - p_e - \delta}{SE} \tag{9}$$

and $$SE = \left[\frac{p_s(1-p_s)}{n_s} + \frac{p_e(1-p_e)}{n_e} \right]^{1/2} \tag{10}$$

where, p_s and p_e are the corresponding observed proportions, and n_s and n_e are the numbers of patients in the two groups.

Even though the theory of hypothesis testing is useful, particularly in planning a clinical trial, the confidence interval approach may be more useful in the analysis, interpretation, and reporting of the accumulated data. Blackwelder states that a hypothesis test tells us whether the observed data are consistent with the null hypothesis and that a confidence interval tells us which hypotheses are consistent with the data.

Blackwelder and Chang (1984). Blackwelder and Chang present the idea of using size graphs in testing whether an experimental therapy is as effective as a standard therapy, but not necessarily more effective. They test

$$H_o : \pi_s \geq \pi_e + \delta \tag{11}$$

$$H_a : \pi_s < \pi_e + \delta \tag{12}$$

where π_s and π_e are success probabilities with standard therapy and experimental therapy, respectively, and δ is the minimum difference of practical interest. If the study is large enough to justify use of the normal approximation to the binomial, the null hypothesis can be tested using the following statistic:

$$z = \frac{(p_s - p_e - \delta)}{s} \tag{13}$$

where,

$$s = \left[\frac{p_s(1-p_s)}{n_s} + \frac{p_e(1-p_e)}{n_e} \right]^{1/2} , \tag{14}$$

p_s and p_e are the observed success proportions with standard and experimental therapy, respectively, and n_s and n_e are the corresponding numbers of patients.

The null hypothesis is rejected at the α-significance level if z is less than the lower $100\alpha\%$ point of the standard normal distribution. The equation below for n is valid in order for the test to have power $1-\beta$ if the two groups are of equal size. The number of patients in each group is then

$$n = (z_{1-\alpha} + z_{1-\beta})^2 [\pi_s(1-\pi_s) + \pi_e(1-\pi_e)] / (\pi_s - \pi_e - \delta)^2 \tag{15}$$

where $z_{1-\alpha}$ and $z_{1-\beta}$ are upper percentage points of the standard normal distribution, n is the number of patients in each group, and π_s and π_e are success probabilities.

Table1

Actual Significance Level of One-Sided Test of H_o: $\pi_s \geq \pi_e + \delta$ at Nominal Significance Level α

n	π_s	π_e	δ	[a]$\alpha = 0.050$	[a]$\alpha = 0.025$	[a]$\alpha = 0.010$
20	0.60	0.50	0.10	0.050	0.028	0.011
20	0.70	0.60	0.10	0.059	0.031	0.014
25	0.80	0.70	0.10	0.056	0.030	0.014
50	0.90	0.80	0.10	0.060	0.030	0.015
100	0.95	0.85	0.10	0.061	0.034	0.016
500	0.99	0.89	0.10	0.060	0.034	0.015
20	0.60	0.40	0.20	0.041	0.021	0.009
20	0.70	0.50	0.20	0.046	0.025	0.010
25	0.80	0.60	0.20	0.054	0.026	0.013
50	0.90	0.70	0.20	0.058	0.031	0.012
100	0.95	0.75	0.20	0.060	0.030	0.014
500	0.99	0.75	0.20	0.057	0.030	0.014

[a]Actual significace level calculated by adding the probabilities, given n, π_s, π_e, δ, and α, of all possible observed outcomes p_s and p_e that yield significant test statistic. Used with permission of Elsevier Science Publishing Inc. (Blackwelder & Chang, 1984).

Hauck and Anderson (1986). The issue of interpreting negative studies occurs frequently with clinical trials. The Freidman, Chalmers, Smith, and Kuebler (1978) approach to interpreting negative studies is referred to as the design-power approach. In this article, the authors focus on a method to interpret and report studies that do not find statistically significant differences. These studies are sometimes referred to as negative studies.

This paper discusses an alternative approach to the design-power approach. The design-power approach determines, irrespective of the observed difference, what differences the study could have been expected to detect. However, there are two important limitations to this method. First, the design-power method does not use the observed

difference. Second, the power calculation can only be an approximation because the actual power depends on the values of the unknown parameters being estimated. The alternative approach to this method examines the similarity of the two treatments given the data. Moreover, the authors discuss some of the methods employed in the analysis of comparative bioavailability studies and show how they can be helpful in the interpretation and reporting of negative studies.

Comparative bioavailability studies involve comparison of drugs that either are equivalent in the active component or are chemical equivalents with respect to the rate and extent of absorption of the drug from its dosage form into systematic circulation. According to Hauck and Anderson, if there are two success probabilities (two means), M_1 and M_2 , equivalence can be expressed as

$$1 - C < \frac{M_1}{M_2} < 1 + C \tag{16}$$

or

$$-\Delta < M_1 - M_2 < \Delta \tag{17}$$

where Δ and C are positive. Situations are restricted to only those estimates of M_1 - M_2 that are normally distributed. Hence,

$$\Delta_1 < M_1 - M_2 < \Delta_2 \quad . \tag{18}$$

The hypotheses for equivalence testing are

$$H_0 : M_1 - M_2 \leq \Delta_1 \text{ or } M_1 - M_2 \geq \Delta_1 \tag{19}$$

$$H_a : \Delta_1 < M_1 - M_2 < \Delta_2 .$$

At this point, it should be noted that the hypotheses are the opposite of those commonly used. For equivalence testing, the alternative hypothesis is that the true difference lies within some specified interval. Westlake (1972) used the confidence interval approach to

equivalence testing. His approach included forming a confidence interval for $M_1 - M_2$ and then concluding equivalence only if that confidence interval falls entirely within the specified interval (Δ_1, Δ_2). Westlake (1981) later discovered that when using the confidence interval to obtain an α-level test, one must use a (100-2α)% confidence interval.

Hauck and Anderson state that the (100-2α)% confidence interval will fall in the equivalence interval only if the null hypothesis is rejected for both of the following tests at the α level:

$$\text{Test A} \qquad H_o: M_1 - M_2 \leq \Delta_1 \qquad H_a: M_1 - M_2 > \Delta_1 \tag{20}$$

$$\text{Test B} \qquad H_o: M_1 - M_2 \geq \Delta_2 \qquad H_a: M_1 - M_2 < \Delta_2 \tag{21}$$

In addition, the confidence intervals must be based on the same test statistic used to test the one-sided hypotheses.

The authors present their approach, the *t*-test method to equivalence testing. The idea of this method is to reject the null hypothesis in favor of the alternative of no more than a 20% difference (assumed interest) if $|T|$ is sufficiently small, where

$$T = \left[\frac{\text{mean difference} - .5(\Delta_1 + \Delta_2)}{\text{standard error}}\right] \tag{22}$$

The *t*-test method should not be used for cases where central *t* approximation has not been verified (indicated by small $\delta = [\frac{1}{2}(\Delta_2 - \Delta_1)/\text{standard error}]$, the estimated non-centrality parameter). In addition, in some cases where $\Delta_1 \neq \Delta_2$, the equivalence curve from the *t*-test is not monotonic because the *t*-statistic is centered around the center of the equivalence interval, not zero.

Wright (1988). Wright focuses on the idea of using prior information when estimating the differences between two proportions. He considers the concept of the amount of information that considers sample sizes in the specific context of estimating the difference between two proportions.

Wright accepts the idea of the following general quote from Berger (1983): "Supplying information is equivalent to removing uncertainty. That is, information supplied = prior uncertainty – posterior uncertainty." In statistical inference, information is generally collected about a parameter Θ using a sample.

Let us assume that two populations are given with respective population proportions, P_1 and P_2. The objective is to estimate

$$P_1 - P_2 . \tag{23}$$

In addition, assume that there exists some prior information concerning P_1 and P_2 which can be expressed by two independent beta prior distributions. The prior distribution for P_i is

$$P_i \sim \frac{\Gamma(\alpha_i + \beta_j)}{\Gamma(\alpha_i)\Gamma(\beta_j)} (P_i)^{\alpha_i - 1}(1 - P_i)^{\beta_i - 1} \tag{24}$$

for $\alpha_i > 0$, $\beta_i > 0$, $0 < P_i < 1$ and $i = 1,2$. If a random sample of size n_i is taken from the i^{th} population and X_i is observed, where X_i is the number of items in the sample from the i^{th} population, then it is obvious that the conditional distribution of X_i given P_i is

$$X_i \Big| P_i \sim \binom{n_i}{x_i} (P_i)^{x_i}(1 - P_i)^{n_i - x_i} \tag{25}$$

for $X_i = 0,1, \ldots, n_i$ and $i = 1,2$. The joint distribution of X_i and P_i is

$$(X_i, P_i) \sim \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \binom{n_i}{x_i} (P_i)^{x_i + \alpha_i - 1}(1 - P_i)^{n_i - x_i + \beta_i - 1} \tag{26}$$

and the marginal distribution of X_i is the beta-binomial distribution

$$X_i \sim \binom{n_i}{X_i} \frac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \frac{\Gamma(X_i+\alpha_i)\Gamma(n_i-X_i+\beta_i)}{\Gamma(n_i+\alpha_i+\beta_i)} \quad . \tag{27}$$

The posterior distribution of P_i given X_i is

$$P_i|X_i \sim \frac{\Gamma(n_i+\alpha_i+\beta_i)}{\Gamma(x_i+\alpha_i)\Gamma(n_i-X_i+\beta_i)} (P_i)^{x_i+\alpha_i-1}(1-P_i)^{n_i-x_i+\beta-1} . \tag{28}$$

The beta distribution is a conjugate prior for the binomial distribution. The Bayes estimator of P_i using loss is

$$L(P_i,\hat{P}_i)=(P_i-\hat{P}_i)^2 \tag{29}$$

where P_i represents the process in estimating P_i given by

$$P_i = E(P_i|X_i)=\frac{x_i+\alpha_i}{n_i+\alpha_i+\beta_i} . \tag{30}$$

To find the optimal choice of n_1 and n_2, for a given n, minimize

$$\sum_{i=1}^{2} \text{var}(P_i|X_i) \tag{31}$$

For more details on equation 31, see Appendix B.

Hauck and Anderson (1992). In this paper, the authors discuss the various types of bioequivalence, their consequences for clinicians and patients, and the status of statistical methods for each type. This paper assumes a single measure of bioavailability from serum levels for oral products and a balanced, two-period crossover design without carryover effects.

In conducting a bioequivalence trial, there are two risks to be controlled. The first is the *consumers' risk*, which is the possibility that an inequivalent formulation would be declared bioequivalent. The second is the *producers' risk*, the possibility that a bioequivalent formulation would fail the criteria to be approved as bioequivalent. Usually, the regulatory authorities control the magnitude of the consumers' risk and the pharmaceutical company decides how much producer's risk they are willing to accept.

Anderson and Hauck (1983) determined that the statement of the statistical hypothesis that should correspond to the objective of a bioequivalence study is as follows: the alternative hypothesis should be the bioequivalence criterion. Hence, the usual type I error would become the consumers' risk and a company could maximize their power of concluding bioequivalence for bioequivalent formulations.

Bioequivalence testing deals mostly with the average levels of the measure of bioavailability. The two formulations are considered average bioequivalent if they are *sufficiently close*.

The hypothesis of average bioequivalence can be stated in terms of either a proportionate or absolute difference. The average bioequivalence hypothesis corresponding to equivalence stated as a proportionate difference is

$$H_a : \Theta_1 \leq \frac{\mu_T}{\mu_R} \leq \Theta_2 \tag{32}$$

where μ_T and μ_R are the average bioavailabilities of the test and reference formulations, respectively. The null hypothesis is

$$H_o : \frac{\mu_T}{\mu_R} < \Theta_1 \text{ or } \frac{\mu_T}{\mu_R} > \Theta_2 . \tag{33}$$

The United States Food and Drug Administration's (FDA) Generic Drug Advisory Committee has recently recommended an interval of $(\Theta_1, \Theta_2) = (.8, 1.25)$ (Hare, 1990).

Average bioequivalence is a special case of *population bioequivalence*, which considers only one aspect of the distribution, namely its mean (or median). Hence, average bioequivalence is an approximation to population equivalence.

Population bioequivalence means a bioequivalence criterion that requires the distribution of the test formulation to be sufficiently similar to that of the reference in some appropriate population (Anderson & Hauck, 1990). According to Anderson & Hauck (1990), population bioequivalence may still be insufficient because there is no mention of the particular individual's similarity of response in the two formulations. Individual bioequivalence means that the bioavailability of the test is sufficiently close to that of the reference for *most* individuals in some appropriate population.

The authors note three approaches to testing individual bioequivalence: (a) formulate a full probabilistic model for the data and estimate all the parameters, followed by some joint test of bioequivalence conditions on those parameters; (b) analyze the individual ratios; and (c) assess each individual to see if they meet a bioequivalence criterion. A tolerance interval and the number of individuals failing that criterion must be sufficiently small.

These three methods all suffer from two problems. First, they are not valid in the presence of period effects. Second, it is difficult to demonstrate that the reference was individual bioequivalent to itself if the within-individual variability is sufficiently large.

Bartolucci and Singh (1993). The authors of this article investigate the idea of therapeutic equivalence testing. Their approach follows the confidence region methodol-

ogy. They define a general class of discrepancy measures between parameters of interest and then apply the Bayesian neighborhood null hypothesis theory to derive posterior confidence regions on those measures. The advantage of the approach is that a variety of beliefs about the behavior of the compounds involved can be accommodated. Moreover, that information can be incorporated into the analysis to demonstrate whether the assumption of therapeutic equivalence is realistic.

Actual biological data was used. The investigators were interested in proving the equivalent failure rate.

This paper is restricted to equivalence in the context of the ratio of survival parameters μ_1 and μ_2 in the exponential family. The null hypothesis is

$$H_o : \eta \leq 1-\Theta \text{ or } \eta \geq 1+\Theta, \tag{34}$$

and the alternative is

$$H_a : 1-\Theta < \eta < 1+\Theta \tag{35}$$

where $\eta = \mu_1 / \mu_2$ and $\Theta = 0.20$. The 1-α posterior probability regions for η can be determined. In addition, it can be determined whether these regions are $\pm\Theta$ or less in width.

The assumption is that n subjects may enter into a trial at random. The data are then separated into a vector of n - r censored data and a vector of r noncensored data. Let

$$T = \sum_{i=1}^{r} t_i + \sum_{k=r+1}^{n} t_k \,'. \tag{36}$$

Then, in the two parameter exponential case for two populations, the joint likelihood function of μ and ω can be denoted by

$$l(\underset{\sim}{\mu},\underset{\sim}{\omega}) = \prod_{j=1}^{J} l(\mu_j,\omega_j) = \prod_{j=1}^{J}\left(\frac{1}{\mu_j}\right)^{r_j} e^{\frac{-1}{\mu_j}(T-n_j\omega_j)}. \tag{37}$$

It is assumed that ω has a prior distribution having density

$$f(\omega,\alpha,\nu)=\frac{\nu e^{\nu\omega}}{\left(e^{\nu\alpha}-1\right)} \tag{38}$$

where $0 < \omega < \alpha$ and $\nu > 0$. The authors make use of a density which is the special limiting case as $\nu \to 0$. From L'Hopital's rule,

$$\lim_{\nu\to 0} f(\omega|\alpha,\nu)=\frac{1}{\alpha}\,. \tag{39}$$

The prior μ is the inverted-gamma-one ($I\,\gamma_1$) family with prior shape parameter, n_0, and prior scale parameter, t_o. This density is as follows:

$$f(\mu|n_o,t_o)=\frac{{t_o}^{n_o}}{\Gamma(n_o)}\left(\frac{1}{\mu}\right)^{n_o+1} e^{\frac{-t_o}{\mu}} \tag{40}$$

where, $0 < \mu$ or $\mu \sim I\gamma_1(n_0, t_0)$.

The supposition is that there are two independent exponential populations. The first population is parameterized by (μ_1, ω_1) and the second by (μ_2, ω_2). It is assumed that there is a sample from the j^{th} population with r_j noncensored observations and $n_j - r_j$ censored data points; j = 1,2. The joint likelihood of the parameters is given by equation 37 with l = 2. The one-to-one transformation is

$$\xi = \mu_1 \qquad \omega_1 = \omega_2$$

$$\eta = \frac{\mu_2}{\mu_1} \qquad \omega_2 = \omega_2$$

The new parameter space is called $\xi = (\xi, \eta, \omega_1, \omega_2)$. Interest is in the test of equivalence of the η under the condition of either $\omega_1 = \omega_2$ or $\omega_1 \neq \omega_2$.

For the condition of equal location parameters, $\omega_1 = \omega_2$, the posterior analysis is performed. The analysis is performed on the induced parameter space from (μ_1, μ_2, ω_1 ω_2) to ($\xi, \eta, \omega_1, \omega_2$). The posterior density of η given the data is as follows:

$$g(\eta|Data, \omega_1 = \omega_2 = \omega) = A * \left[\frac{(B-C)}{(D-E)}\right] * F * G \tag{41}$$

where

$$A = \frac{\eta^{-(r_2+n_0-1)}}{(n_1 + n_2\ \eta)^3}\Gamma(r_1 + r_2 + 2n_0 - 1)$$

$$B = \left[\left(\frac{T_2 + t_o}{\eta}\right) + T_1 + t_o\right]^{r_1+r_2+2n_0-1}$$

$$C = \left[\left(\frac{T_2 + t_o}{\eta}\right) + T_1 + t_o - \delta\right]^{r_1+r_2+2n_0-1}$$

$$D = \left[\left(\frac{T_2 + t_o}{\eta}\right) + T_1 + t_0\right]^{r_1+r_2+2n_0-1}$$

$$E = \left[\left(\frac{T_2 + t_o}{\eta}\right) + T_1 + t_0 - \delta\right]^{r_1+r_2+2n_o-1}$$

$$F = \left[\prod_{j=1}^{2}\Gamma(r_j + n_0)\right]^{-1}$$

$$G = \int_0^{\delta}\ \prod_{j=1}^{2}(T_j + t_0 - n_{jx})^{-(r_j+n_0)}\ dx$$

where δ= min (observations, a). Equation 41 can be used to compute the posterior probability that η lies in some equivalence region of interest $(\varepsilon_1,\varepsilon_2)$.

For the unequal location parameters, $\omega_1 \neq \omega_2$, the posterior density of η given the data is

$$g(\eta|Data,\omega_1 \neq \omega_2)=\frac{NUM}{DEN} \qquad (42)$$

where

$$A(\omega_1,\omega_2)=\frac{C}{D}$$

$$C=\eta^{-(n_0+r_2+1)}$$

$$NUM=\int_0^{\delta_1}\int_0^{\delta_2} A(\omega_1,\omega_2)d\omega_1 d\omega_2$$

$$D=\left[(T_1+t_0n_1\omega_1)+\frac{1}{\eta(T_2+t_0-n_2\omega_2)}\right]^{-(2n_0+r_1+r_2-2)}$$

$$DEN=\int_0^{\infty} NUM \quad d\eta$$

δ_j= min (observations, a_j) j = 1,2.

The posterior probability that η lies in some equivalence region of interest can also be computed for this case.

Pham-Gia and Turkkan (1993). In this article, the authors derive the expression for the posterior distribution of p_1- p_2. The result that the natural conjugate property of the beta family for binomial sampling is one of the basic results in Bayesian analysis. For example, if p has a beta(α,β) prior then the posterior distribution of p is beta($\alpha + x$, $\beta + n - x$), where x is the number of successes and n is the total number of observations in

the sample phase. There is a lack of results for the Bayesian treatment of the density of p_1- p_2. This paper establishes the precise expression of that density and then applies it to compute the posterior distribution.

The connection between the beta (and its generalizations) and the hypergeometric function permits the authors to obtain the desired density. Even though there are four Appell hypergeometric functions in two variables, only F_1 and F_3 were used.

Definition: Let a_1, a_2, b_1, b_1, and c be real or complex numbers with c different from a negative integer. Appell's first hypergeometric function in two variables, F_1, is defined by:

$$F_1(a_1,b_1,b_2;c;x_1,x_2)=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(a_1,m+n)}{(c,m+n)}(b_1,m)(b_2,n)\frac{x_1^{\,m}}{m!}\frac{x_2^{\,n}}{n!}$$

where

$$(a,m)=a(a+1)..(a+m-1)=\Gamma(a+m)/\Gamma(a),m>0, \text{ with } (a,0)=1.$$

Similarly, F_3 is defined by

$$F_3(a_1,a_2,b_1,b_2;c;x_1,x_2)=\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{(a_1,m)(a_2,n)}{(c,m+n)}(b_1,m)(b_2,n)\frac{x_1^{\,m}}{m!}\frac{x_2^{\,n}}{n!}.$$

It is established that F_1 and F_3 converge for $|x_1|<1$ and $|x_2|<1$. F_1 and F_3 are related by the following:

$$F_1(a,b_1,b_2;c;x,y)=(1-y)^{-b_2}F_3(a,c-a,b_1;b_2;c;x,y/(y-1)) \quad (43)$$

$$=(1-x)^{-b_1}F_3(c-a,a,b_1,b_2;c;x/(x-1),y) \quad (44)$$

Theorem: Let $p_i \sim$ beta(α_i, β_i), i =1,2, be independent variables. Then $p=p_1$- p_2 has the following density:

For $0<p\le 1$, $f(p)=$

$$B(\alpha_2,\beta_1)p^{\beta_1+\beta_2-1}(1-p)^{\alpha_2+\beta_1-1}F_1(\beta_1,\alpha_1+\beta_1+\alpha_2+\beta_2-2,1-\alpha;$$

$$\beta_1+\alpha_2;1-p,1-p^2)/A \quad (45)$$

and for $-1 \le p < 0$

$$f(p) = B(\alpha_1, \beta_2)(-p)^{\beta_1+\beta_2-1}(1+p)^{\alpha_1+\beta_2-1} F_1(\beta_2, 1-\alpha_2,$$

$$\alpha_1+\alpha_2+\beta_1+\beta_2-2; \alpha_1+\beta_2; 1-p^2, 1+p)/A, \quad (46)$$

where $A = B(\alpha_1, \beta_1)B(\alpha_2, \beta_2)$

Moreover, if $\alpha_1+\alpha_2 > 1$ and $\beta_1+\beta_2 > 1$, we have

$$f(0) = B(\alpha_1+\alpha_2-1, \beta_1+\beta_2-1)/A. \quad (47)$$

For the summary of the proof to this theorem, see Appendix C.

The density of $p_1 - p_2$ will have a variety of shapes since beta(α, β) can have a wide variety of shapes depending on the values of α and β. The posterior distributions of p_1 and p_2 are independent beta(α_1+x_1, $\beta_1+n_1-x_1$) and beta(α_2+x_2, $\beta_2+n_2-x_2$) when independent sampling of p_1 and p_2 gives x_1 and x_2 favorable outcomes out of n_1 and n_2 observations, respectively.

The posterior distribution of p is given by Equations (45), (46), and (47) with $\alpha_i + x_i$ replacing α_i and $\beta_i + n_i - x_i$ replacing β_i, i = 1,2. The distribution with its density given by Equations (45), (46) and (47) is called a beta-difference distribution. This distribution (beta-difference distribution) is closed under independent dual Bernoulli sampling.

Gopalan and Berry (1998). This article considers the problem of multiple comparisons from a Bayesian point of view. Multiple comparisons are special types of multiplicities. Bayesian multiple comparison procedures enable direct probability calculations of hypotheses of equality and inequality among treatment means. The following are

examples of multiplicities: (a) multiple comparisons (treatments), (b) meta-analysis (studies), and (c) multicenter trials (centers).

The authors make use of Dirichlet process priors (DPPS) to develop a methodology for obtaining posterior probabilities of hypotheses under two different prior likelihood combinations. Gibbs sampling is used for evaluation because the solution is analytically intractable.

CHAPTER 3

THERAPEUTIC EQUIVALENCE

Equivalence studies are often needed to develop better tolerated therapies after effective treatments have been identified (Durrleman & Simon, 1990). Therapeutic equivalence refers to a new treatment being as effective as a standard treatment. The question of equivalence arises in the treatment of infectious diseases, cancer, and other illnesses when one is considering a new therapy that is thought to be as effective as but perhaps not more effective than, an existing therapy (Blackwelder, 1982).

Two alternative formulations of the same drug are said to be bioequivalent when equal amounts of the formulations produce equal therapeutic effects (Kirkwood, 1981). Two different drugs or formulations of the same drug are called bioequivalent if they are absorbed into the blood and become available at the drug action site at about the same rate and concentration (Berger & Hsu, 1996).

In order to give a more direct definition of therapeutic equivalence, some preliminary definitions from the Food and Drug Administration (FDA) will be beneficial. The starting point for understanding therapeutic equivalence is the term *therapeutic agent* or, as it is usually called, *therapeutic moiety*. This term refers to the substance in a drug product that actually achieves the intended effect in the diagnosis cure, mitigation, treatment, or prevention of disease or in affecting the structure or function of the human body. Although different substances may produce the same ultimate therapeutic effect, they are not necessarily identical therapeutic agents. For example, various narcotics produce

analgesia, but do so through different, although related, therapeutic moieties. On the other hand, the same therapeutic moiety may appear in slightly different chemical forms (e.g., as different salts or esters of the same molecule). To distinguish these separate forms, the term, *active drug ingredient* is used; each salt or ester of a therapeutic agent is a unique active drug ingredient. For example, tetracycline hydrochloride and tetracycline phosphate complex are distinct active ingredients containing the same therapeutic moiety (FDA, 1979).

Drug product means a finished dosage (e. g., tablet, capsule, or solution) that contains the active drug ingredient, generally, but not necessarily, in association with inactive ingredients.

Bioavailability means the rate and extent to which the active ingredient or active moiety is absorbed from a drug product and becomes available at the site of action. Bio-availability may be assessed by measurements intended to reflect the rate and extent to which the active ingredient or active moiety becomes available at the site of action.

Pharmaceutical equivalents means drug products that contain identical amounts of the identical active drug ingredient (i.e., the same salt or ester of the same therapeutic moiety) in identical dosage forms, but not necessarily containing the same inactive ingredients. The drug products also meet the identical compendial or other applicable standard of identity, strength, quality and purity, including potency and, where applicable, content uniformity, disintegration times and/or dissolution rates.

Pharmaceutical alternatives means drug products that contain the identical therapeutic moiety, or its precursor, but not necessarily in the same amount or dosage form, or as the same salt or ester. Each such drug product individually meets either the identical or its own respective compendial or other applicable standard of identity,

strength, quality, and purity, including potency, and where applicable, content uniformity, disintegration times, and/or dissolution rates.

Bioequivalence means the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study. Where there is an intentional difference in rate (e.g., in certain controlled release dosage forms), certain pharmaceutical equivalents or alternatives may be considered bioequivalent if there is no significant difference in the extent to which the active ingredient or moiety from each product becomes available at the site of drug action. This applies only to the differences in the rate at which the active ingredient or moiety becomes available at the site of drug action. This is intentional, but is not essential to the attainment of effective body drug concentrations on chronic use. It is usually reflected in the proposed labeling
and is considered medically insignificant for the drug (Nation & Samson 1994; FDA, 1992).

According to the FDA, drug products may be evaluated as therapeutically equivalent if (a) they are pharmaceutical equivalents in that they contain identical amounts of the same dosage form, and they meet identical compendial or other applicable standards of identity, strength, quality, and purity; (b) they are bioequivalent in that either they present no known or potential problem, or if they are shown to meet an appropriate bioequivalence standard (bioequivalence refers to the comparative rates and

extents of absorption of drug products into the human body); (c) they are adequately labeled; and (d) they are manufactured in compliance with current good manufacturing practice (FDA, 1979).

According to Blanchard and Sawchuck (1979), bioavailablity is the rate and extent of absorption of the drug from its dosage form into systematic circulation. According to Hauck and Anderson (1992), two drugs are considered bioequivalent if their bioavailabilities differ by less than some meaningful limit.

Therapeutic equivalence is important in many arenas: for example, the government, the pharmaceutical industry, and the medical industry. Showing therapeutic equivalence is worthwhile because it may show that a new drug is less toxic, easier to administer, or less expensive than the established therapy. In addition, a generic drug may be proven to be therapeutic equivalent to a competing product. This would allow the generic drug company to compete with the originator's product (Huque et al., 1990).

CHAPTER 4

ACTIVE CONTROL TRIALS

An active control trial is a clinical trial, sometimes referred to as a positive control trial. An active control trial is similar to the placebo control with the exception that there is no placebo. A placebo is an inactive substance (with respect to the experiment under investigation) that has all of the likenesses of the drug(s) under investigation: for instance, taste, shape, appearance.

Active control trials usually involve two therapies, an experimental and a reference. The therapies are such that they have similar taste, appearance, shape, etc. These similarities allow the therapies to be distributed such that the patient is "blind" to the treatment received and the physician is "blind" to the treatment prescribed (Modell & Houde, 1958). The technique mentioned above is known as the double-blind technique.

According to Modell and Houde (1958), nine forces influence data in clinical evaluations: pharmacodynamic actions, dosage, choice of subject, use of controls, collection of data, sensitivity of the method, placebo actions, sensitivity of the method, placebo actions, bias, and forces extraneous to the experiment. The regimen is another factor that should be considered when designing a clinical trial. The fundamental assumptions of the active control is that the active control drug would have performed better than a placebo, had a placebo been used in the trial (Makuch & Johnson ,1989).

There are many ethical and practical reasons for active control trials' popularity. An ethical reason for their popularity often stated in the literature is that it is unethical to deny a patient immediate access to a known effective treatment (Leber, 1986). Some of the practical reasons include smaller sample sizes, subject recruitment enhancement, and lower dropout rates.

Active control trials are criticized for the absence of information regarding the reference therapy's ability to perform better than a placebo, the increased possibility of misleading conclusions by misinterpretations of the data, and the inability to distinguish between true drug effects and improvements attributable to a placebo effect.

CHAPTER 5

PROPOSED METHODOLOGY

Introduction

This chapter focuses on the proposed methodology for comparing proportions. It begins with the comparison of two proportions by examining their ratio. These ideas were then extended to three proportions as presented in Integration for Comparing Three Proportions, below. Being more specific, that section focuses entirely on the integration for comparing three proportions. This is appropriate because integrating three proportions invoked some dimensionality problems.

In the latter sections of this chapter, we discuss weighted likelihood ratio. The ratio was described by I. J. Good (1950) as the "factor in favor of the hypothesis, H, in virtue of the result of the experiment."

Methodology

In this chapter, methodology and notation are developed in order to fulfill the research objectives. The development begins with the comparison of two proportions, p_1 and p_2. We are interested in testing

$$H_o: \eta \leq 1-\theta \qquad (48)$$

$$H_a: 1-\theta < \eta < 1+\theta$$

where $\eta = p_1 / p_2$ and $\theta = 0.2$.

The posterior density of the ratio of two proportions is given by

$$p(\eta_1|data) = \frac{\int g(\xi,\eta_1)l(\xi,\eta_1)d\xi}{\int\int g(\xi,\eta_1)l(\xi,\eta_1)d\xi d\eta_1} \quad (49)$$

where $g(\xi,\eta_1)$ denotes the joint prior density function, $l(\xi,\eta_1)$ denotes the joint likelihood function, $\eta_1 = p_1/p_2$, and $\xi = p_2$. The reader should refer to Appendix D for the details of the derivations concerning Equation (49). Confidence regions were then determined. A basic problem when assessing equivalence is specification of an acceptable difference. The FDA Guidelines specified 20% for comparative bioavailability studies. In other circumstances, such a specified limit may not be as clear-cut. If we let $p(\Delta)$ represent the *p*-value associated with an equivalence test using limits $\pm\Delta$ [or $p(C)$ for limits $1 \pm C$], we can then examine a plot of $p(\Delta)$ against Δ or $p(C)$ against C. See Appendices E and F for the results of the confidence regions for the ratio of two proportions for various values of α and β. These calculations were done using MathCad software version 8. MathCad is a standard calculation software by MathSoft, Inc.

The methodology for the ratio of two proportions was then extended to three proportions. This extension is discussed in the subsequent section.

Integration for Comparison of Three Proportions

Integration issues. Applications having the multiplicities are among the most difficult faced by statisticians and other researchers. Unfortunately, this happens frequently in Bayesian analysis. Multiple comparisons often invoke analytically intractable solutions. Oftentimes, integrands do not behave well in certain regions. One must pursue numerical methods of integration when the multiple integration is nontractable. The

analyst uses the numerical methods of multiple integration to change the integrand such that it is computable within the range of interest. Even though there are several problems that may occur when computing multiplicities, one must remember that the overall most important task of the Bayesian analyst is to complete the statistical inference process.

Method of integration. Determining the posterior density for the three proportions

$$p(\eta_1,\eta_2|data)=\frac{\int g(\xi,\eta_1,\eta_2)l(\xi,\eta_1,\eta_2)d\xi}{\iiint g(\xi,\eta_1,\eta_2)l(\xi,\eta_1,\eta_2)d\xi d\eta_1 d\eta_2} \tag{50}$$

presented many challenges. The formula for the posterior is as follows: where

$$\eta_1=\frac{p_1}{p_2} \qquad \eta_2=\frac{p_1}{p_3} \qquad \xi=p_1.$$

The numerator was calculated directly. On the other hand the denominator had to be integrated using a change of variables. Let

$$\eta_1=\frac{p_1}{p_2}, \qquad \eta_2=\frac{p_1}{p_3}, \qquad \xi=p_1,$$

then

$$p_2=\frac{\xi}{\eta_1} \quad \text{and} \quad p_3=\frac{\xi}{\eta_2}.$$

The absolute value of the Jacobian is

$$|J|=\begin{vmatrix} 0 & \frac{-\xi}{\eta_1^2} & 0 \\ 0 & 0 & \frac{-\xi}{\eta_2^2} \\ 1 & \frac{1}{\eta_1} & \frac{1}{\eta_2} \end{vmatrix}=\frac{\xi^2}{\eta_1^2\eta_2^2}.$$

The prior is

$$\pi(p_i) = \frac{1}{B((\alpha,\beta))} p_i^{\alpha-1}(1-p_i)^{\beta-1}, \text{ and the likelihood is } l(p_i) = p_i^{y_i}(1-p_i)^{n_i-y_i}.$$

The joint prior is as follows:

$$g(\xi,\eta_1,\eta_2) = \left(\frac{1}{B(\alpha,\beta)}\right)^3 (\xi)^{\alpha-1}(1-\xi)^{\beta-1}\left(\frac{\xi}{\eta_1}\right)^{\alpha-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta-1}\left(\frac{\xi}{\eta_2}\right)^{\alpha-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta-1}\frac{\xi^2}{\eta_1^2\eta_2^2}.$$

The joint likelihood is as follows:

$$l(\xi,\eta_1,\eta_2) = (\xi)^{y_1}(1-\xi)^{n_1-y_1}\left(\frac{\xi}{\eta_1}\right)^{y2}\left(1-\frac{\xi}{\eta_1}\right)^{n_2-y_2}\left(\frac{\xi}{\eta_2}\right)^{y_3}\left(1-\frac{\xi}{\eta_2}\right)^{n_3-y_3}.$$

So the triple integral is now,

$$\iiint g(\xi,\eta_1,\eta_2)l(\xi,\eta_1,\eta_2)d\xi d\eta_1 d\eta_2,$$

which is the integral of the following:

$$\left(\frac{1}{B(\alpha,\beta)}\right)^3 (\xi)^{\alpha+y_1-1}(1-\xi)^{\beta+n_1-y_1-1}\left(\frac{\xi}{\eta_1}\right)^{\alpha+y_2-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta+n_2-y_2-1}\left(\frac{\xi}{\eta_2}\right)^{\alpha+y_3-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta+n-y_3-1}\frac{\xi^2}{\eta_1^2\eta_2^2}$$

The calculation of this integral using the data is shown in the section titled Three Proportions (Chapter 6).

A separate procedure was done to obtain $\eta_3=p_3/p_2$. This procedure yielded $P(\eta_3|\text{data})$. To find the density of

$$\eta_3 = \frac{p_3}{p_2}, \text{ let } T = \frac{\eta_1}{\eta_2} \text{ where } T \text{ is } \eta_3. \text{ So, } T = \frac{\eta_1}{\eta_2} \quad X = \eta_2$$

then $TX = \eta_1$.

The absolute value of the Jacobian is

$$|J| = \left\| \begin{bmatrix} x & 0 \\ 1 & 1 \end{bmatrix} \right\| = X.$$

Assuming η_1 and η_2 are independent, we get the following:

$$f(\eta_1, \eta_2) = \int p(\eta_1, \eta_2 | data) d\eta_1 \int p(\eta_1, \eta_2 | data) d\eta_2$$

and use substitution to get f(T, X). Hence, $f(T) = \int f(T, X) dx$.

Derivation of equivalence cube axes. This section shows the derivation of the "equivalence cube" axes. Assume $p_1 > p_2$; then, according to equivalence testing as discussed by Hauck and Anderson (1986), $p_1 - p_2 \leq \Delta$, where $\Delta = 0.2$. So,

$$p_1 - p_2 \leq 0.2$$

$$(p_1 / p_2) - 1 \leq (0.2 / p_2)$$

$$(p_1 / p_2) \leq 1 + (0.2 / p_2)$$

Hence, for both sides

$$1 - (0.2 / p_2) \leq (p_1 / p_2) \leq 1 + (0.2 / p_2). \quad (51)$$

In addition, assume $p_1 > p_3$. So,

$$p_1 - p_3 \leq 0.2$$

$$(p_1 / p_3) - 1 \leq (0.2 / p_3)$$

$$(p_1 / p_3) \leq 1 + (0.2 / p_3)$$

Hence, for both sides

$$1 - (0.2 / p_3) \leq (p_1 / p_3) \leq 1 + (0.2 / p_3). \quad (52)$$

Moreover, assume $p_3 > p_2$. So,

$$p_3 - p_2 \leq 0.2$$

$$(p_3 / p_2) - 1 \leq (0.2/ p_2)$$

$$(p_3 / p_2) \leq 1 + (0.2/ p_2)$$

Hence, for both sides

$$1-(0.2/ p_2) \leq (p_3 / p_2) \leq 1 + (0.2/ p_2). \quad (53)$$

Weighted Likelihood Ratio

According to Thomas Jefferson, "mathematical reasoning and deduction are a fine preparation for investigating the abstruse speculations of the law (Osteyee & Good, 1974)." Hence, the purpose of this section is to explain the process of weighing evidence.

Let D represent the data from an experiment modeled by a distribution function, Φ, with parameter vector $\theta \in E^{\vartheta}$. [E^{ϑ} denotes the Euclidean ϑ–space.] The probability density function or the probability mass function of $\Phi(D|\theta)$ can be denoted by $\phi(D|\theta)$. Denote the likelihood function of θ depending on D by

$$\ell(\theta) \propto \phi(D|\theta). \quad (54)$$

If one is interested in testing the hypothesis, (H), that θ belongs to a Borel subset, $\theta \in H \subset E^{\vartheta}$, against the alternative that $\theta \in H^c$, where $H^c \cap H = \varnothing$, then consider the information that concludes this section. The prior odds for H are represented by

$$O(H) = \frac{p(H)}{p(\overline{H})}. \quad (55)$$

The posterior probability of H by Bayes theorem is as follows:

$$P\left(H\middle|\underset{\sim}{D}\right)=\frac{\psi\left(\underset{\sim}{D}\middle|H\right)P(H)}{\psi\left(\underset{\sim}{D}\middle|H\right)P(H)+\psi\left(\underset{\sim}{D}\middle|\overline{H}\right)P\left(\overline{H}\right)} \tag{56}$$

where, $\psi(D \mid K)$ is the Stieltjes integral of appropriate dimensionality. K may by H or H^c. Furthermore, the posterior odds for H is then

$$O\left(H\middle|\underset{\sim}{D}\right)=\frac{P\left(H\middle|\underset{\sim}{D}\right)}{P\left(\overline{H}\middle|\underset{\sim}{D}\right)}=\frac{\psi\left(\underset{\sim}{D}\middle|H\right)}{\psi\left(\underset{\sim}{D}\middle|\overline{H}\right)}*\frac{P(H)}{P\left(\overline{H}\right)}=\mathrm{L}_{\underset{\sim}{D}}*\mathrm{O(H)} \tag{57}$$

where

$$\mathrm{L}_{\underset{\sim}{D}}(\mathrm{H})=\frac{\psi\left(\underset{\sim}{\mathrm{D}}\middle|H\right)}{\psi\left(\underset{\sim}{D}\middle|\overline{H}\right)}. \tag{58}$$

Equation 58 is referred to as the Bayes' factor by Good (1950). A formal definition is as follows: Let H_1 and H_2 be two competing hypotheses related to some evidence B. H_1 and H_2 may be thought of as "events" both in the a priori probability spaces (Ω, H_1, p) and (Ω, H_2, P_B), where P_B is the conditional probability given B. Then the weight of evidence in favor of H_1 as opposed to H_2, provided by B, may be defined as

$$W(H_1/H_2 : B)=\log\frac{O(H_1\ H_2|B)}{O(H_1\ H_2)},$$

where $O(H_1 / H_2 \mid B)$ is the odds in favor of H_1 as opposed to H_2 given B, and $O(H_1/H_2)$ is the odds in favor of H_1 as opposed to H_2. Weight of evidence may be positive or negative. For the detailed derivation of the weighted likelihood ratio, see Appendix G.

CHAPTER 6

APPLICATION OF THE METHODOLOGY

Problem Statement

Most phase III studies are designed with several endpoints in mind, one of which is to compare treatments with respect to rate of response, toxicity, or both. Analyses of two data sets from a published breast cancer study and lung cancer study, respectively, were performed.

The objective of the trial related to the first data set was to compare CHOP (cytoxan, adriamycin, vincristine, and prednisone) to BCOP (BCNU, cytoxan, vincristine, and prednisone) in diffuse histiocytic non-Hodgkin's lymphoma (Hayward et al., 1977). The objective of the trial related to the second data set was to compare CAMF (cyclophosphamide, adriamycin, methotrexate with folinic acid) with CAP (cyclophoshamide, adriamycin, cis-platinum) and CA (cyclophosphamide, adriamycin) in an advanced non-small-cell carcinoma of the lung with respect to rate of response and toxicity (Berkson, 1958). The third data set was formed to use as a means of comparison to the second data set.

With respect to the context of this writing, the analyses performed are focused on the response aspect of the objectives. The ratio of the proportions were examined with each data set.

Data Set Information

Two Proportions. Hayward et al. (1977) give well-defined response criteria for metastatic breast cancer. CHOP was compared to BCOP in diffuse histiocytic non-Hodgkin's lymphoma in a Southeastern Cancer Study Group (SECSG) protocol.

There were 53 patients randomized and evaluated on BCOP and 59 patients randomized and evaluated on CHOP. The complete responses were considered. There were 19 complete responses (CR) on BCOP and 33 CR on CHOP.

Three proportions. Patients were randomized to one of three treatments in an advanced non-small-cell carcinoma of the lung trial. Patients were appropriately stratified. The treatments were as follows: (a) CAMF (cyclophosphamide, adriamycin, methotrexate with folinic acid, (b) CAP (cycloplosphamide, adriamycin, cis-platinum), and (c) CA (cyclophosphamide, adriamycin). The total number of observations involved in the analyses was 339. The three treatments were to be compared with respect to their ability to achieve a CR or partial response (PR). There were 13 responses out of 98 possibilities for treatment CAMF. The treatment CAP had a response proportion of 9 out of 113. There were 4 responses out of 128 for CA.

The third data set, a test set, consists of three treatments. The three treatments were PA, which had a response of 13 out of 20; CT with 12 out of 20 responses; and ON, which had 11 out of 20 responses.

Analysis and Results

Two proportions. Regarding the comparison of two proportions, the posterior density was determined using the beta-binomial distribution in Bayesian methodology.

After deriving the posterior distribution for η_1, where $\eta_1 = p_1/p_2$, $E(\eta|D)$ and $E\{[\eta - E(\eta|D)]^2|D\}$ were computed for prior values of α and β. $Z_{0.05}$ is the multiplier needed for computing the 90% limits, based on asymptotic normal distribution theory. Consequently, the 90% limits are calculated and compared to a predetermined region width of 0.20. The selection of this specified width is from the bioavailability guideline produced by the FDA.

The results of the calculations of the confidence region using the sample data and Bayes' estimates for the ratio of the two proportions can be seen in Appendices E and F. Based on the previously mentioned guidelines, we can conclude conclude that a necessary condition for equivalence has not been established.

Three proportions. Determining an efficient method to compare three proportions pairwise, we first derived the posterior density using a beta-binomial prior. This density, given in Method of Integration (Chapter 5), as well as the weighted likelihood ratio derivation given in Appendix G, contained a triple integral, which was analytically intractable; however, the Monte Carlo was used to alleviate this dimensionality problem.

The triple integrals involved in this research were calculated using MathCad. A Fortran program was also developed to do the calculations. This program was useful because it allowed a means for comparing its solutions to those of MathCad. In addition, the Fortran program provided numerical results that could be plotted to actually observe the very peculiar behavior of the functions. For the benefit of the reader, the numerical results from this program are provided in Appendix H. Plots of the numerical results can be viewed in Appendices I-K. Following is the Monte Carlo procedure ap-

plied to the triple integral part of the density using the first data set from the section titled Three Proportions, with three proportions.

Originally, the triple integral was

$$\int_{.43}^{1.57}\int_{.75}^{1.25}\int_{0}^{1}\xi^{\alpha+y1-1}(1-\xi)^{\beta+n1-y1}\left(\frac{\xi}{\eta_1}\right)^{\alpha+y2-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta+n2-y2}\left(\frac{\xi}{\eta_2}\right)^{\alpha+y3-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta+n3-y3-1}$$

$$\frac{\xi^2}{\eta_1^2\eta_2^2}d\xi d\eta_1 d\eta_2.$$

Because of necessary constraints, we have the following integral:

$$\int_{.43}^{1.57}\int_{.75}^{1.25}\int_{0}^{1}\xi^{\alpha+y1-1}(1-\xi)^{\beta+n1-y1}\left(\frac{\xi}{\eta_1}\right)^{\alpha+y2-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta+n2-y2}\left(\frac{\xi}{\eta_2}\right)^{\alpha+y3-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta+n3-y3-1}$$

$$\frac{\xi^2}{\eta_1^2\eta_2^2}d\xi d\eta_1 d\eta_2,$$

but if we let $\omega = \xi/\eta_1$; then, when $\eta_1 = 1, \omega = \xi$ and $\eta_1 = 1.25$, $\omega = \xi/1.25$ so $d\omega = (-\xi/\eta_1^2)$ $d\eta_1$. Similarly, we let $z = \xi/\eta_2$; then, when $\xi/\eta_2 = 1$, $z = \xi$ and $\eta_2 = 1.57$, $z = \xi/1.57$ so $dz = -(-\xi/\eta_2^2)$.

Hence, the triple integral becomes

$$\int_0^1 \frac{\xi^{\alpha+y1-1}(1-\xi)^{\beta+n1-y1}}{B(\alpha,\beta)}\cdot F_1(\xi)\cdot F_2(\xi)d\xi \qquad (59)$$

where

$$F_1(\xi) = \frac{B(\alpha+y2,\beta+n2-y2)}{B(\alpha,\beta)}\left[IB\left(\xi|\alpha+y2,\beta+n2-y2\right) - IB\left(\frac{\xi}{1.57}|\alpha+y2,\beta+n2-y2\right)\right]$$

$$F_2(\xi) = \frac{B(\alpha+y3,\beta+n3-y3)}{B(\alpha,\beta)}\left[IB\left(\xi|\alpha+y2,\beta+n2-y2\right) - IB\left(\frac{\xi}{1.57}|\alpha+y2,\beta+n2-y2\right)\right]$$

For the purpose of identification within Appendix H, let

$$F_0(\xi) = \int_0^1 \frac{\xi^{\alpha+y1-1}(1-\xi)^{\beta+n1-y1}}{B(\alpha,\beta)}\, d\xi.$$

Next, the confidence regions were determined. These regions using the sample data and Bayes' estimates are given in Appendices L and M, respectively. The pairwise comparison was then made by seeing if the posterior of the ratio endpoints fell within the *equivalence cube* as shown in Appendices N and O. Based on the results of the cubes, we see that a necessary condition of equivalence has not been established. From the third data set, we see that a necessary condition of equivalence has been established (i.e., the endpoints of the confidence regions fell within the equivalence region). For numerical results for this data set see Appendix P. The equivalence cubes for the third data set can be seen in Appendices, Q and R.

Weighted likelihood ratio (WLR). The WLR was calculated for all combinations of η_1, η_2, and η_3 in the null hypotheses. A test data set with values expected to be equivalent was used to calculate another group of WLRs for all combinations of η_1, η_2, and η_3 in the null hypotheses. All of these WLRs were calculated using Bayes' estimates. This data set was derived for the total purpose of comparing the results to the real data set The results of these calculations can be seen in Appendices S, T, U,V, and W.

We can see from the results that the values of the WLR for the cancer data were extremely small. This is further evidence that the proportions are not equivalent. On the other hand, the values of the WLR for the test data were greater than one for each pairwise combination. This suggests that there is sufficient evidence for equivalence.

Discussion

These findings support the hypothesis of equivalence for the scale and nonscale parameters. As a result, a researcher would conclude that a necessary condition has not been established for equivalence between the three cancer therapies regarding their effectiveness in the treatment of the disease. However, a researcher would conclude that a necessary condition has been established for equivalence between the three test therapies that were formed for a means of comparison. The results of the data analysis by both methods, the equivalence region and the WLR, indicate the proper response of equivalence using this methodology. It is also important to note that the sample data estimates and the Bayes' estimates gave similar results. The findings using this methodology are consistent with the findings of the classical methods.

CHAPTER 7

SUMMARY AND POSSIBLE EXTENSIONS

Introduction

In this chapter we briefly summarize the methodology developed in Chapters 5 and 6 and propose extensions of this methodology.

Summary

In this paper, we considered the problem of assessing therapeutic equivalence of three independent proportions. The methodology used is based on the Bayesian methodology of Pham-Gia and Turkkan (1993); however, we considered the ratio of proportions instead of the difference of two proportions. Even more challenging, we studied three proportions instead of two. The WLR was also incorporated into this work. Chapters 5 and 6, along with the appendices, actually show how this methodology was implemented.

Possible Extensions

Main areas of this research relate to equivalence, the integration involved in determining the posterior densities, inference constructions, and the basis for establishing a sufficient condition for equivalence.

This research assumed that all p_i's were independent. It would be of great interest to extend this research with the assumption that the p_i's are not independent. One pos-

sible method to be explored for this dependence assumption is the concept expressed in the De Finetti theorem on exchangeable variables. This theorem is as follows: To every infinite sequence of exchangeable random variables (X_n) having values in {0,1}, there corresponds a probability distribution F concentrated on [0,1] such that

$$P\{ X_1 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, Xn = 0\}$$

$$= \int_0^1 \Theta^k (1-\Theta)^{n-k} F(d\Theta) \quad \text{for all n and } 0 \le k \le n.$$

The distribution F may be regarded as the prior for the random parameter Θ (Heath & Sudderth, 1976). Exchangeable is defined as follows: The random variables, $X_1, \ldots, X_n$, are exchangeable if the n! permutations,

$$(X_{k_1}, \ldots, X_{k_n})$$

have the same *n*-dimensional probability distribution (Freedman & Diaconis, 1980).

Chapter 5 presented some issues pertaining to the integration involved in deriving the posterior density for the pairwise comparison of three proportions. There is definitely a need for the development of more methods that would be beneficial in handling multi-dimensional integration problems from a computational perspective. Moreover, the triple integral in this research involved some beta functions with very interesting behavior. It would be of interest to direct attention to the study of the behavior of such complicated functions.

Calculations from this research were done using the Monte Carlo method as well as basic integration principles. But other methods, such as the Gibbs sampling algorithm, need more exploration.

This research may appear to be directed toward the Bayesian statistician; however, other areas of statistics and mathematics would benefit greatly from further research

in this area. For example, even though the subject of power is inconsistent in our Bayesian framework, it would be of interest to see how it would be applicable in this research. Chapter 2 discussed some related research. Hauck and Anderson (1992) presented types of bioequivalence and some related considerations. Their work has two main areas where further research is needed. First, statisticians need methods for assessing population bioequivalence and methods for individual bioequivalence. Second, there is a need for more methods that are appropriate for measures of bioavailability.

In this section, possible extensions, the encouragement of the development of other statistical inference constructions should be included because of the diversity of trials regarding the manner in which data is collected and because the amount of information available before, during, and after the trial needs to be handled. In addition, other statistical inference constructions to the problem of equivalence bring new ideas that complement previous ones and help others unfold.

LIST OF REFERENCES

Anderson S., & Hauck, W. (1983). A new procedure for testing equivalence in comparative bioavailability and other clinical trials. *Communications in Statistical Theory and Methods, 12, 2663-2692.*

Anderson S. & Hauck, W. (1990). Consideration of individual bioequivalence. *Journal of Pharmacokinectics and Biopharmaceutics, 18,* 259-273.

Bartolucci, A. A., & Dickey, J. M. (1977). Comparative Bayesian and traditional inference for gamma modeled survival data. *Biometrics, 33,* 343-354.

Bartolucci, A. A., Katholi, C. R., & Birch, R. (1992). Interim analysis of failure time data - A Bayesian approach. *Environmetrics, 3,* 465-477.

Bartolucci, A. A., & Singh, K. P. (1993). Therapeutic equivalence - A Bayesian approach. *Proceedings of the Biopharmaceutical Section of the American Statistical Association.* 349-357.

Berger, R. L. & Hsu, J. C. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statistical Science, 11,* 283-319.

Berger, T. (1983). "Information theory and coding theory." In Kotz, S., Johnson, N. L., and Read, C. B.(Eds.), *Encyclopedia of Statistical Sciences 4* (pp. 124-141). New York, NY.

Berkson, J. (1958). Smoking and lung cancer: some observations on two recent reports. *Journal of the American Statistical Association, 53,* 28-38.

Birch, R., & Bartolucci, A. A. (1983). Determination of the hyperparameters of a prior probability model in survival analysis. *Computer Programs in Biomedicine, 17,* 89-94.

Blackwelder, W. C. (1982). "Proving the null hypothesis" in clinical trials. *Controlled Clinical Trials, 3,* 345-353.

Blackwelder, W. C., & Chang, M. A. (1984). Sample size graphs for "proving the null hypothesis." *Controlled Clinical Trials, 5,* 97-105.

Blanchard, J., & Sawchuck, R. J. (1979). Drug bioavailability: an overview. In Blanchard, J., Sawchuck, R. J. & Brodie, B. B. (Eds.), *Principles and Perspectives in Drug Bioavailability,* (pp.1-19).

Carey, D., Bartolucci, A. A., & Singh, K. P. (1998, December 30-31). *Assessing therapeutic equivalence (TE) in clinical trials using the beta-binomial distribution.* University of Alabama at Birmingham, Birmingham, Alabama. Paper presented at the International Statistical Application Conference, India. To be published.

Dickey, J. M. (1979). Beliefs about beliefs, a theory for statistical assessments of subjective probabilities. In J. Beruardo (Ed.)., *In Bayesian Statistics.* Spain: University of Valencia Press.

Durrleman, S., & Simon, R. (1990). Planning and monitoring of equivalence studies. *Biometrics, 48,* 329-336.

FDA (Food and Drug Administration (1979)). Therapeutically equivalent drugs. availability of list. *Federal Register 44,* 2932-2953.

FDA (1992) Food and Drug Administration bioavailability and bioequivalence requirement *Federal Register 57,* 17007-18001.

Fluehler, H., Grive, A. P., Mandallaz, D., Mau, J., & Moser, H.A. (1983). Bayesian approach to bioequivalence assessment: an example. *Journal of Pharmaceutical Sciences, 72,* 1178-1181.

Freedman, D., & Diaconis, P. (1980) Finite exchangeable sequences. *The Annals of Probability 8,* 745-764.

Freidman, J. A., Chalmers, T. C., Smith, J., & Kuebler, R. R. (1978). The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. *New England Journal of Medicine, 299,* 690-694.

Glatstein, E., & Makuch, R. W. (1984). Illusion and reality: practical pitfalls in interpreting clinical trials. *Journal of Clinical Oncology,* 2, 488-497.

Good, I. J. (1950). *Probability and the weighing of evidence.* London: Charles Griffin.

Gopalan, R., & Berry, D., (1998). Bayesian multiple comparisons using Dirichlet process priors. *Journal of the American Statistical Association, 93,* 1130-1139.

Hare, D. (1990). The orange book: the Food and Drug Administration's advice on on therapeutic equivalence. *American Pharmaceutical,* NS30: 35-37.

Hauck, W. W., & Anderson, S. (1986). A proposal for interpreting and reporting negative studies. *Statistics in Medicine, 5,* 203-209.

Hauck, W. W., & Anderson, S. (1992). Types of bioequivalence and related statistical considerations. *International Journal of Clinical Pharmacology, Therapy and Toxicology, 30,* 181-187.

Hayward, J. L., Carbone, P.P., Heuson, J.C., Kimaoka, S., Segaloff, A., & Rubens, R.D. (1977). Assessment of response to therapy in advanced breast cancer. *Cancer*, 39, 1289-94.

Heath, D., & Sudderth, W. (1976). De Finetti's theorem on exchangeable variables. *The American Statistician, 30*, 188-189.

Huque, M., Dubey, S., & Fredd, S. (1990). Establishing therapeutic equivalence with clinical endpoints. *Proceedings of the Biopharmaceutical Section of the American Statistical Association*. 46-52.

Jefferson, C., Bartolucci, A. A., & Singh, K. P. (1997). Survival–bioequivalence: empirical Bayesian approach with non-conventional Iinference framework. *Proceedings of the International Congress on Modeling and Simulation. Hobart, Tasmania, Australia,3*, 1023-1028.

Kirkwood, T. (1981). Bioequivalence testing–a need to rethink. *Biometrics, 37*, 589-594.

Leber, P. (1986). The placebo control in clinical trials (A view from the FDA). *Psychopharmacology Bulletin*, 22, 30-32.

Makuch, R. & Johnson, M.(1989). Issues in planning and interpreting active control equivalence studies. *Journal of Clinical Epidemiology, 42*, 503-511.

Metzler, C. M. (1974). Bioavailability–a problem in equivalence, *Biometrics, 30*, 309-317.

Nation, R. L. & Samson, L. N. (1994). Bioequivalence requirements for generic products. *Pharmacology and Therapeutics, 62*, 41-55.

Modell, W., & Houde, R. W (1958). Factors influencing clinical evaluation of drugs. *Journal of the American Medical Association, 167*, 2190-2199.

Osteyee, D., & Good, I. J. (1974). Information, Weight of Evidence, the Singularity between Probability Measures and Signal Detection. *Lecture Notes in Mathematics*. New York: Springer-Verlag.

Pham-Gia, T., & Turkkan, N. (1993). Bayesian analysis of difference of two proportions. *Communications in Statistics–Theory and Methods, 12*, 1755-1771.

Selwyn, M.R., Dempster, A.P., & Hall, N. R. (1981). A Bayesian approach to bioequivalence for the 2x2 changeover design. *Biometrics, 37*, 261-269.

Westlake, W. (1972). Use of confidence intervals in analysis of comparative bioavailability trials, *Journal of Pharmaceutical Science, 61*, 1340-1341.

Westlake, W. (1981). Response to bioequivalence testing–a need to rethink. *Biometrics 37*, 591-593.

Wright, T. (1988). **A note on optimal allocation and an ordering of the family of beta prior distributions.** *Communications in Statistical Theory and Methods, 17,* 123-138.

APPENDIX A

WEIGHTED AVERAGE

$$\sum_r p_r f_r = \sum \frac{P(H_r|H)P(E|H_r)}{P(E|\overline{H})}$$

$$= \sum \frac{P(H_r|H)P(E|H_r * H)}{P(E|\overline{H})}$$

$$= \sum \frac{P(E * H_r|H)}{P(E|\overline{H})} = \frac{P(E * H|H)}{P(E|\overline{H})} = \frac{P(E|H)}{P(E|\overline{H})}.$$

APPENDIX B

OPTIMAL CHOICE OF n_1 AND n_2

To find the optimal choice of n_1 and n_2, for given n_1, we minimize

$$\sum_{i=1}^{2} \text{var}(P_i|X_i) = \sum_{i=1}^{2} \frac{(X_i+\alpha_i)(n_i - X_i + \beta_i)}{(n_i + X_i + \beta_i)^2 (n_i - X_i + \beta_i + 1)}$$

which is the expected value of the loss $[(P_1 - P_2) - (P_1^* - P_2^*)]^2$

with respect to the posterior distributions in equation 28, subject to the constraint

$$n = \sum_{i=1}^{2} n_i .$$

$\sum_{i=1}^{2} \text{var}(P_i|X_i)$ depends on X_1 and X_2, initially unknown.

Average $\sum_{i=1}^{2} \text{var}(P_i|X_i)$ with respect to the marginal distributions given in equation 27.

Hence, we minimize, $E(x_1, x_2)\left|\sum_{i=1}^{2} \text{var}(P_i|X_i)\right|$ subject to the constraint.

X_i is a beta-binomial random variable with parameters n_i, α_i and β_i. Observe

$$E_{x_i}(X_i) = \frac{n_i \alpha_i}{\alpha_i + \beta_i}$$

$$\text{var}(X_i) = E_{x_i}(X_i^2) - [E_{x_i}(X_i)]^2 = \frac{n_i \alpha_i \beta_i (n_i + \alpha_i + \beta_i)}{(\alpha_i + \beta_i)^2 (\alpha_i + \beta_i + 1)}$$

$$\text{So, } E_{(x_1,x_2)}\left[\sum_{i=1}^{2} \text{var}(P_i|X_i)\right] = \sum_{i=1}^{2} \frac{E_{x_i}[(x_i+\alpha_i)(n_i - x_i + \beta_i)]}{(n_i + \alpha_i + \beta_i)^2 (n_i + \alpha_i + \beta_i + 1)}$$

$$= \sum_{i=1}^{2} \frac{\alpha_i \beta_i}{(\alpha_i + \beta_i + 1)(n_i + \alpha_i + \beta_i)(\alpha_i + \beta_i)}$$

Using the constraint, $n_1 + n_2 = n$, we obtain the following optimal allocation:

$$n_i^* = \frac{n + \sum_{j=1}^{2} (\alpha_j + \beta_j)}{\sum_{j=1}^{2} \left(\frac{\alpha_j \beta_j}{(\alpha_j + \beta_j + 1)(\alpha_j + \beta_j)} \right)^{1/2}} \left(\frac{\alpha_i \beta_i}{(\alpha_i + \beta_i + 1)(\alpha_i + \beta_i)} \right)^{1/2} - (\alpha_i + \beta_i) \text{ for i} = 1,2.$$

APPENDIX C

SUMMARY OF PROOF TO THEOREM

Proof (summary): Since f(p) is the convolution of f_1 and f_2, $f = f_1 * f_2$, for $0 < p \le 1$ we have

$$f(p) = \int_0^{1-p} (p+v)^{\alpha_1-1}(1-p-v)^{\beta_1-1} v^{\alpha_2-1}(1-v)^{\beta_2-1} dv / A \qquad (60)$$

$$= p^{\alpha_1-1}(1-p)^{\beta_1-1} \int_0^{1-p} v^{\alpha_2-1}(1-v)^{\beta_2-1}(1+v/(1-p))^{\beta_1-1} dv / A,$$

where $A = B(\alpha_1,\beta_1)\, B(\alpha_2,\beta_2)$.

Changing the integration variable to w= v/(1-p) and applying Picard's theorem*, we obtain

$$f(p) = p^{\alpha_1-1}(1-p)^{\alpha_2+\beta_1-1} B(\alpha_2,\beta_1) F_1(\alpha_2, 1-\beta_2, 1-\alpha_1; \beta_1+\alpha_2; 1-p, 1-1/p) \ / A. \quad (61)$$

Applying $F_1(a,b_1,b_2;c;x,y) = (1-x)^{c-(a+b_1)}(1-y)^{-b_2}$ to change the variables inside F_1 we obtain equation 45 where, $|1 - p| < 1$ and $|1 - p^2| < 1$, as required for the convergence of F_1. Equation 46 can be proved in a similar way. For p = 0 Equation (60) above reduces to equation 47 if $\alpha_1 + \alpha_2 > 1$ and $\beta_1 + \beta_2 > 1$. Keep in mind that f(0) can be undefined in other cases.

*Picard's theorem: Let a, b_1,b_2, and c be real or complex numbers. If Re(a) and Re(c-a) are positive and $F_1(a,b_1,b_2;c;x_1,x_2)$ converges, then

$$F_1(a,b_1,b_2;c;x_1,x_2) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \int_0^1 u^{a-1}(1-u)^{c-a-1}(1-ux_1)^{-b_1}(1-ux_2)^{-b_2} du$$

Pham-Gia, T., & Turkkan, N. (1993). Bayesian analysis of difference of two proportions. *Communications in Statistics - Theory and Methods, 12*, 1755-1771

APPENDIX D

POSTERIOR DENSITY OF THE RATIO OF TWO PROPORTIONS

Let p_i represent the proportions, where i = 1,2. The prior density is given by

$$\pi(p_i) = \frac{1}{\beta(\alpha,\beta)}(p_i)^{\alpha-1}(1-p_i)^{\beta-1}$$

where, $\alpha > 0$, $\beta > 0$, and $0 \le p_i \le 1$.

Furthermore, the likelihood function is given by

$$l(p_i) = p_i^{y_i}(1-p_i)^{n_i-y_i}$$

where, i = 1,2.

To derive the density of the ratio of proportion 1 and proportion 2, we first obtain the joint prior density and joint likelihood functions. Assuming independence, the joint prior is as follows:

$$\pi(p) = \frac{1}{\beta(\alpha,\beta)}(p_1)^{\alpha-1}(1-p_1)^{\beta-1}\frac{1}{\beta(\alpha,\beta)}(p_2)^{\alpha-1}(1-p_2)^{\beta-1}$$

$$= \left(\frac{1}{\beta(\alpha,\beta)}\right)^2 (p_1)^{\alpha-1}(1-p_1)^{\beta-1}(p_2)^{\alpha-1}(1-p_2)^{\beta-1}$$

In addition, the joint likelihood is

$$\Pi l(p) = p_1^{y_1}(1-p_1)^{n_1-y_1}p_2^{y_2}(1-p_2)^{n_2-y_2}$$

The transformation is as follows:

Let

$$\eta_1 = \frac{p_1}{p_2} \qquad \xi = p_2$$

then

$$\eta_1\xi = p_1.$$

The absolute value of the Jacobian is

$$|J| = \left\| \begin{bmatrix} \frac{\delta p_1}{\delta \xi} & \frac{\delta p_1}{\delta \eta_1} \\ \frac{\delta p_2}{\delta \xi} & \frac{\delta p_2}{\delta \eta_1} \end{bmatrix} \right\| = \left\| \begin{bmatrix} \eta_1 & \xi \\ 1 & 0 \end{bmatrix} \right\| = |-\xi| = \xi$$

Now, rewriting p_1 and p_2, we have

$$g(\xi, \eta_1) = \left(\frac{1}{\beta(\alpha, \beta)} \right)^2 (\eta_1 \xi)^{\alpha-1} (1 - \eta_1 \xi)^{\beta-1} (\xi)^{\alpha-1} (1 - \xi)^{\beta-1} \xi$$

and

$$l(\xi, \eta_1) = (\eta_1 \xi)^{y_1} (1 - \eta_1 \xi)^{n_1 - y_1} (\xi)^{y_2} (1 - \xi)^{n_2 - y_2}.$$

So, the posterior density is

$$p(\eta_1 | data) = \frac{\int_0^1 \left(\frac{1}{\beta(\alpha, \beta)} \right)^2 (\eta_1 \xi)^{\alpha + y_1 - 1} (1 - \eta_1 \xi)^{\beta + n_1 - y_1 - 1} (\xi)^{\alpha + y_2 + 1 - 1} (1 - \xi)^{\beta + n_2 - y_2 - 1} d\xi}{\int_{1-\hat{p}_2}^{1} \int_0^1 \left(\frac{1}{\beta(\alpha, \beta)} \right)^2 (\eta_1 \xi)^{\alpha + y_1 - 1} (1 - \eta_1 \xi)^{\beta + n_1 - y_1 - 1} (\xi)^{\alpha + y_2 + 1 - 1} (1 - \xi)^{\beta + n_2 - y_2 - 1} d\xi d\eta_1}.$$

APPENDIX E

CONFIDENCE REGIONS FOR TWO PROPORTIONS (USING SAMPLE DATA)

The following data are used in Table E1:

Treatment	*CR's*	*Non – CR's*	*Total*	$\hat{p}_i$'s i = 1,2
CHOP	33	26	59	0.559
BCOP	19	34	53	0.358
TOTAL	52	60	112	

Using the sample data, p_1 / p_2 is as follows:

$$\frac{\hat{p}_1}{\hat{p}_2} = 1.561.$$

Table E1

Confidence Regions for Ratio of Two Proportions (α=.1)

Lower endpoint	Upper endpoint
1.098	2.022
1.099	2.021
1.96	2.024

For these calculations, p_1 is the proportion of CHOP and p_2 is the proportion of BCOP. At the (α=.1) level, we can see that both of the confidence endpoints are not within the equivalence region of (.441, 1.558).

APPENDIX F

CONFIDENCE REGIONS FOR TWO PROPORTIONS (USING BAYES ESTIMATES)

The following data are used in Table F1:

The Bayes estimate for various combinations of α and β is as follows:

α	β	$\frac{\hat{p}_1}{\hat{p}_2}$
2	3	1.51
3	2	1.48
2	4	1.51
2	2	1.51

Table F1

Confidence Regions for Ratio of Two Proportions (α =. 1)

Prior Parameters			
α	β	Lower endpoint	Upper endpoint
2	3	1.048	1.972
3	2	1.019	1.941
2	3	1.046	1.974
2	2	1.045	1.969

For these calculations, p_1 is the proportion of CHOP and p_2 is the proportion of BCOP. At the (α=.1) level, we can see that both of the confidence endpoints are not within the equivalence region of (.441, 1.558).

APPENDIX G

WEIGHTED LIKELIHOOD RATIO DERIVATION

We begin with the following hypothesis:

$H_o : \eta_1 = \eta_2 = 1$

H_a : **at least one distinct**

The formula of the weighted likelihood ratio is

$$L_D(H) = \frac{\int l(\xi, \eta_H) g(\xi | \eta_H) d\xi}{\iiint l(\xi, \eta) g(\xi, \eta) d\xi d\eta}.$$

We let

$$\eta_1 = \frac{p_1}{p_2}, \quad \eta_2 = \frac{p_1}{p_3}, \text{and} \quad \xi = p_1$$

which implies that

$$p_2 = \frac{\xi}{\eta_1} \quad \text{and } p_3 = \frac{\xi}{\eta_2}.$$

The absolute value of the Jacobian is

$$|J| = \begin{vmatrix} 0 & \frac{-\xi}{\eta_1^2} & 0 \\ 0 & 0 & \frac{-\xi}{\eta_2^2} \\ 1 & \frac{1}{\eta_1} & \frac{1}{\eta_2} \end{vmatrix} = \frac{\xi^2}{\eta_1^2 \eta_2^2}.$$

The joint prior is as follows:

$$g(\xi, \eta_1, \eta_2) = \left(\frac{1}{B(\alpha, \beta)}\right)^3 (\xi)^{\alpha-1}(1-\xi)^{\beta-1}\left(\frac{\xi}{\eta_1}\right)^{\alpha-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta-1}\left(\frac{\xi}{\eta_2}\right)^{\alpha-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta-1}\frac{\xi^2}{\eta_1^2\eta_2^2}.$$

The joint likelihood is as follows:

$$l(\xi, \eta_1, \eta_2) = (\xi)^{y_1}(1-\xi)^{n_1-y_1}\left(\frac{\xi}{\eta_1}\right)^{y2}\left(1-\frac{\xi}{\eta_1}\right)^{n_2-y_2}\left(\frac{\xi}{\eta_2}\right)^{y_3}\left(1-\frac{\xi}{\eta_2}\right)^{n_3-y_3}.$$

It is known that

$$g(\xi|\eta_H) = \frac{g(\xi,\eta)}{g(\eta)},$$

and substituting in the values of the null gives

$$g(\xi|\eta_H) = \frac{\xi^{3\alpha-1}(1-\xi)^{3\beta-2-1}}{B(3\alpha,3\beta-2)}$$

and $l(\xi,\eta_1,\eta_2)g(\xi|\eta_H) = \left(\frac{1}{B(3\alpha,3\beta-2)}\right) B(3\alpha+y_1+y_2+y_3, n_1-y_1+n_2-y_2+n_3-y_3)$

+ (3β-2). Also,

$$\int\int\int l(\xi,\eta_1,\eta_2) g(\xi,\eta_1,\eta_2)\, d\xi d\eta_1 d\eta_2$$

which is equal to the triple integral of the following terms :

$$\left(\frac{1}{B(\alpha,\beta)}\right)^3 (\xi)^{\alpha+y_1-1}(1-\xi)^{\beta+n_1-y_1-1}\left(\frac{\xi}{\eta_1}\right)^{\alpha+y_2-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta+n_2-y_2-1}\left(\frac{\xi}{\eta_2}\right)^{\alpha+y_3-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta+n-y_3-1}.$$

$$\frac{\xi^2}{\eta_1^{\,2}\eta_2^{\,2}}$$

$$\text{Let } Q = \int\int\int (\xi)^{\alpha+y_1-1}(1-\xi)^{\beta+n_1-y_1-1}\left(1-\frac{\xi}{\eta_1}\right)^{\beta+n_2-y_2-1}\left(1-\frac{\xi}{\eta_2}\right)^{\beta+n_3-y_3-1}\left(\frac{\xi}{\eta_1}\right)^{\alpha+y_2-1}\left(\frac{\xi}{\eta_1}\right)^{\alpha+y_3-1}$$

$$\frac{\xi^2}{\eta_1^{\,2}\eta_2^{\,2}} d\xi d\eta_1 d\eta_2$$

then

$$L_D(H) = \frac{B(3\alpha+y_1+y_2+y_3, n_1-y_1+n_2-y_2+n_3-y_3+3\beta-2)(B(\alpha,\beta))^3}{B(3\alpha,3\beta-2)Q}$$

For the hypothesis

$H_0 : \eta_3 = 1$ against $H_a : \eta_3 \neq 1$

$\eta_3 = \frac{p_3}{p_2}$ we get the following weighted likelihood ratio:

$$L_D(H) = \frac{\frac{\beta(y_3 + y_2 + 2\alpha, n_2 - y_2 + 2\beta - 1 + n_3 - y_3)}{B(2\alpha, 2\beta - 1)}}{\int\int \left(\frac{1}{B(\alpha,\beta)}\right)^3 (\xi)^{\alpha+y_2+1-1}(1-\xi)^{n_2-y_2+\beta-1}(\eta_3\xi)^{\alpha+y_3-1}(1-\eta_3\xi)^{n_3-y_3+\beta-1} d\xi d\eta_3}.$$

For the following hypothesis:

$H_o : \eta_1 = \eta_3 = 1$

H_a : at least one distinct,

the formula of the weighted likelihood ratio is

$$L_D(H) = \frac{\frac{B(3\alpha + y_1 + y_2 + y_3, n_1 - y_1 + n_2 - y_2 + n_3 - y_3 + 3\beta - 2)}{B(3\alpha, 3\beta - 2)}}{\frac{S}{(B(\alpha,\beta))^3}}$$

where

$$S = \int\int\int (\eta_1\xi)^{\alpha+y_1-1}(1-\eta_1\xi)^{\beta+n_1-y_1-1}(\xi)^{\alpha+y_2-1}(1-\xi)^{n_2-y_2+\beta-1}(\eta_3\xi)^{\alpha+y_3-1}(1-\eta_3\xi)^{n_3-y_3+\beta-1} d\eta_3$$

$\xi^2 d\xi d\eta_1$

$\eta_1 = \frac{p_1}{p_2}, \eta_3 = \frac{p_3}{p_2}$ and $\xi = p_2$.

For the following hypothesis:

H_o : $\eta_2 = \eta_3 = 1$

H_a : at least one distinct,

The formula of the wlr is

$$L_{\underset{\sim}{D}}(H) = \frac{\dfrac{\mathrm{B}(3\alpha + y_1 + y_2 + y_3, n_1 - y_1 + n_2 - y_2 + n_3 - y_3 + 3\beta - 2)}{\mathrm{B}(3\alpha, 3\beta - 2)}}{\dfrac{G}{(\mathrm{B}(\alpha,\beta))^3}}$$

where

$$G = \int\int\int (\eta_2\xi)^{\alpha+y_1-1}(1-\eta_2\xi)^{\beta+n_1-y_1-1}(\xi)^{\alpha+y_3-1}(1-\xi)^{n_3-y_3+\beta-1}\left(\frac{\xi}{\eta_3}\right)^{\alpha+y_2-1}\left(1-\frac{\xi}{\eta_3}\right)^{n_2-y_2+\beta-1}$$

$$\frac{\xi^2}{\eta_3^{\,2}} d\xi d\eta_2 d\eta_3$$

$\eta_2 = \frac{p_1}{p_3}, \eta_3 = \frac{p_3}{p_2}$ and $\xi = \mathrm{p}_3$.

APPENDIX H

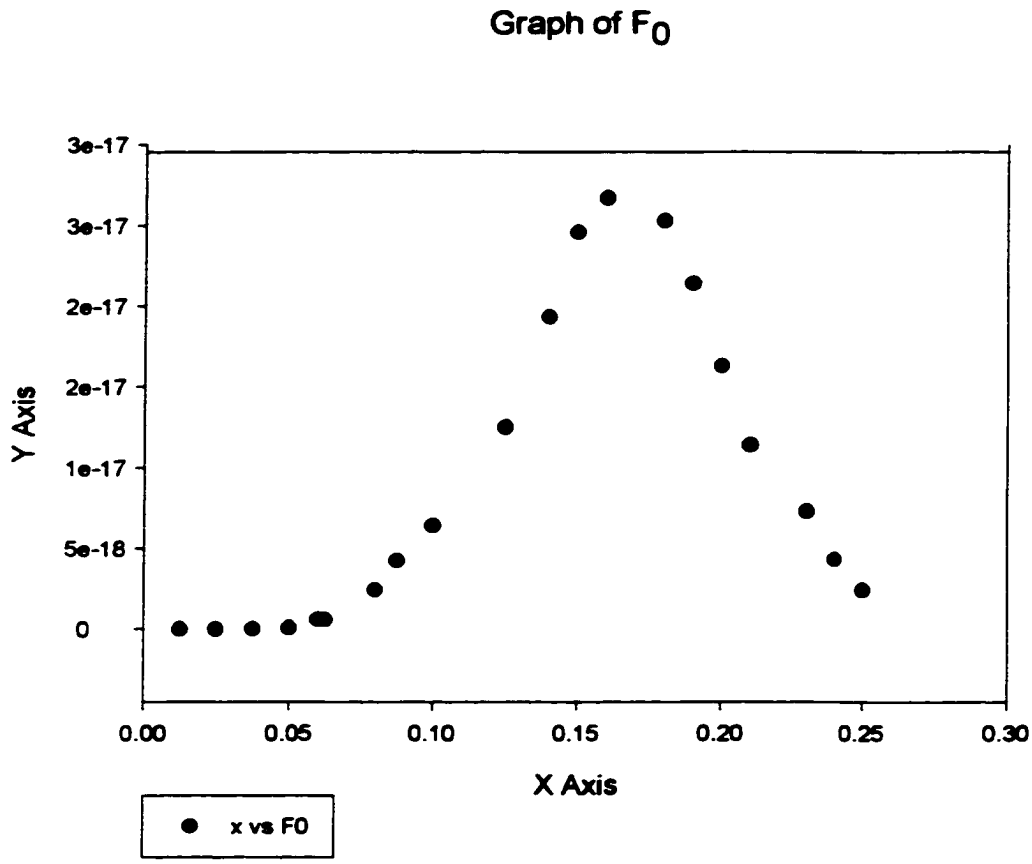
NUMERICAL RESULTS FROM PROGRAM

The triple integrals to be evaluated involved in obtaining the necessary results for this research were divided into the product of three functions, F_0, F_1,and F_2 . See Analysis and Results for the defined functions.

X	F_1	F_2	F_0
1.2500000D-02	3.5118048D-38	1.9070339D-10	9.1337580D-27
2.5000000D-02	2.0565025D-35	1.0735979D-09	4.9401818D-23
3.7500000D-02	5.0386122D-34	2.5451022D-09	4.6933530D-21
5.0000000D-02	3.3484250D-33	4.2299302D-09	8.4472426D-20
6.2500000D-02	1.0844325D-33	5.7825989D-09	6.0673987D-19
7.5000000D-02	2.2223883D-32	6.9823433D-09	2.4230908D-18
8.7500000D-02	3.3132116D-32	5.6020567D-04	6.4202797D-18
1.0000000D-01	1.2573367D-15	5.6020598D-04	1.2539481D-17
1.1250000D-01	1.2573367D-15	5.6020590D-04	1.9318083D-17
1.2500000D-01	0.0000000D+00	5.6020552D-04	2.4582273D-17
1.3750000D-01	0.0000000D+00	7.0051598D-09	2.6696913D-17
1.5000000D-01	0.0000000D+00	6.2992687D-09	2.5345765D-17
1.6250000D-01	0.0000000D+00	5.5400915D-09	2.1419401D-17
1.7500000D-01	0.0000000D+00	4.7803391D-09	1.6338322D-17
1.8750000D-01	0.0000000D+00	4.0568839D-09	1.1371879D-17
2.0000000D-01	0.0000000D+00	3.3930040D-09	7.2850083D-18
2.1250000D-01	0.0000000D+00	2.8011638D-09	4.3251907D-18
2.2500000D-01	0.0000000D+00	2.2857758D-09	2.3932310D-18
2.3750000D-01	0.0000000D+00	1.8456468D-09	1.2397629D-18
2.5000000D-01	0.0000000D+00	1.4759844D-09	6.0349167D-19

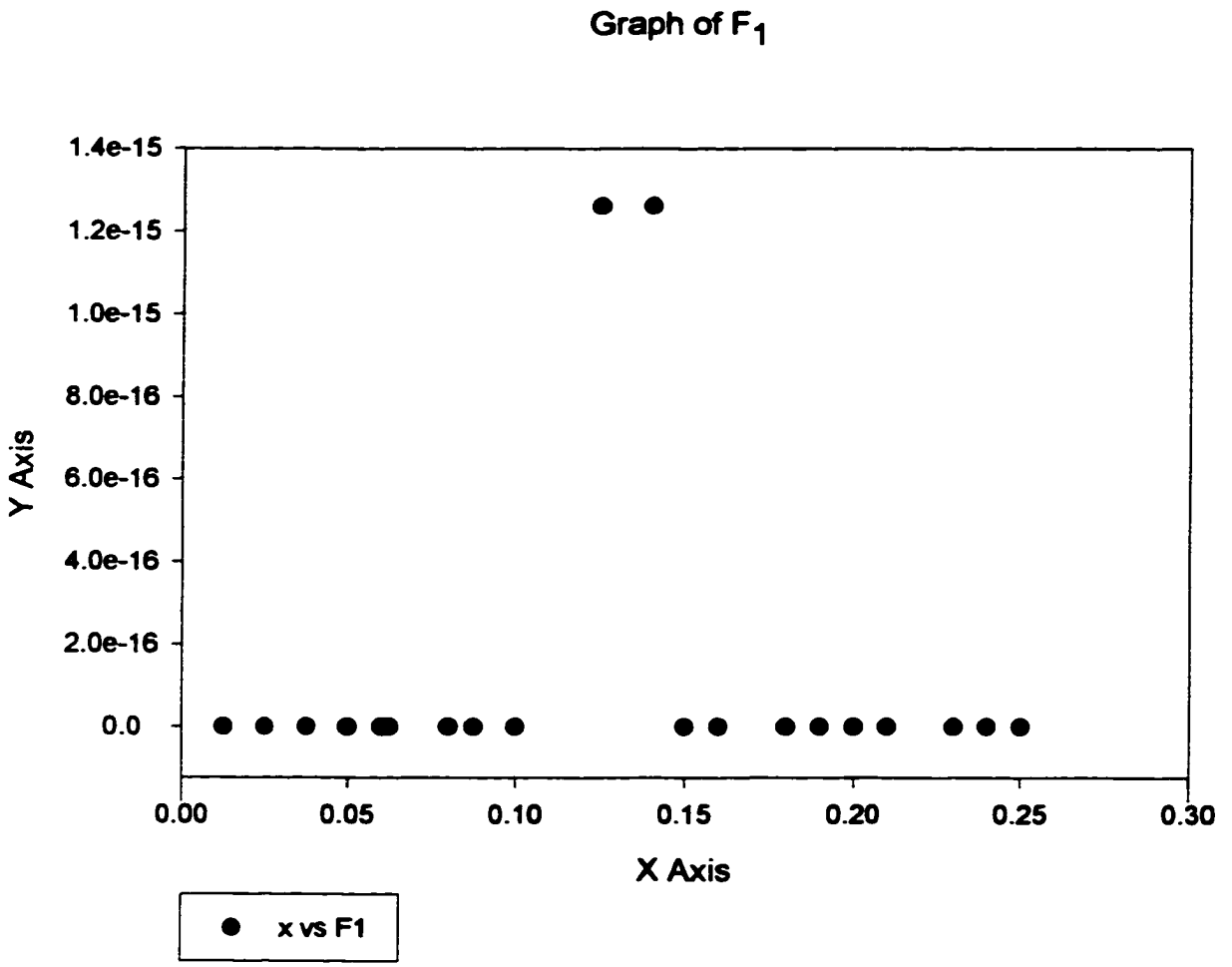
APPENDIX I

GRAPH OF F_0



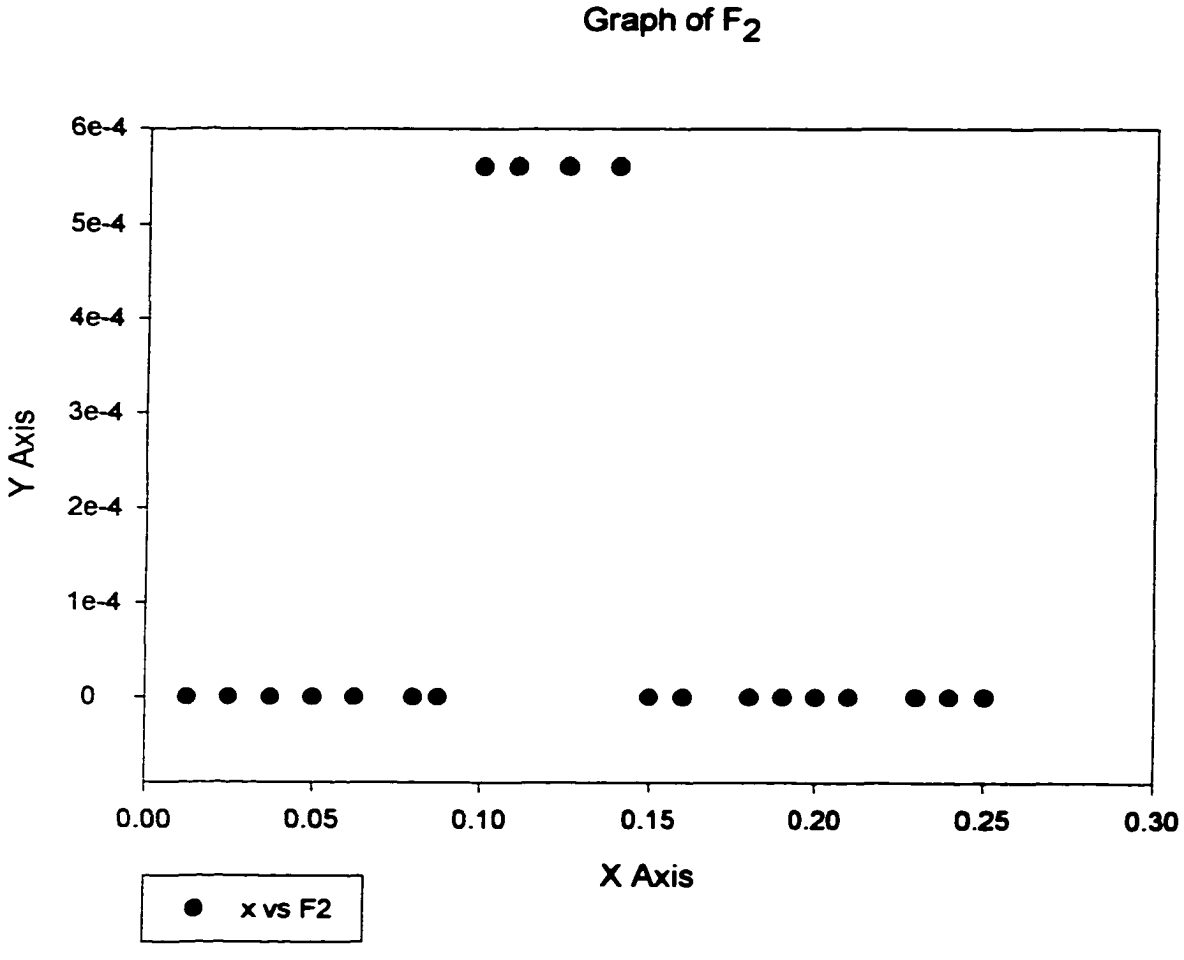
APPENDIX J

GRAPH OF F_1



APPENDIX K

GRAPH OF F_2



APPENDIX L

CONFIDENCE REGIONS FOR PAIRWISE COMPARISON OF THREE PROPORTIONS (USING SAMPLE DATA)

The following data are used in Table L1 below:

	CR or PR(%)	No Response(%)	Total
CAMF	13 (13.3%)	85 (86.7%)	98
CAP	9 (8%)	104 (92%)	113
CA	4 (3%)	124 (97%)	128
Total	26	313	339

The proportions from the sample data are as follows:

$\eta_1 = p_1 / p_2 = 1.67$

$\eta_2 = p_1 / p_3 = 4.24$

$p_3 / p_2 = .39$

where p_1 is the proportion of CAMF, p_2 is the proportion of CAP and p_3 is the proportion of CA. The p's are estimated values.

Table L1

Pairwise Comparison of Three Proportions (α=.1)
Confidence Regions for the Posterior of the Ratio of Proportions Using Sample Data

p_1 / p_2	p_1 / p_3	p_3 / p_2
(1.554,1.785)	(4.009,4.471)	(0.271,0.509)
(1.555,1.785)	(4.007,4.473)	(0.270,0.510)

APPENDIX M

CONFIDENCE REGIONS FOR PAIRWISE COMPARISON OF THREE PROPORTIONS (USING BAYES' ESTIMATES)

The following data are used in Table M1 results below:

	CR or PR(%)	No Response(%)	Total
CAMF	13 (13.3%)	85 (86.7%)	98
CAP	9 (8%)	104 (92%)	113
CA	4 (3%)	124 (97%)	128
Total	26	313	339

The Bayes estimates are as follows:

2	3	1.56	1.796	.87
3	2	1.52	1.44	1.06

where p_1 is the proportion of CAMF, p_2 is the proportion of CAP and p_3 is the proportion of CA.

Table M1

Pairwise Comparison of Three Proportions (α=.1)
Confidence Region for the Posterior of the Ratio of Proportions Using Bayes' Estimates

Prior parameters				
α	β	p_1/p_2	p_1/p_3	p_3/p_2
2	3	(1.445,1.675)	(1.565,2.027)	(0.751,0.989)
3	2	(1.405,1.635)	(1.207,1.673)	(0.940,1.179)

APPENDIX N

EQUIVALENCE CUBE FOR $\alpha = 2, \beta = 3$

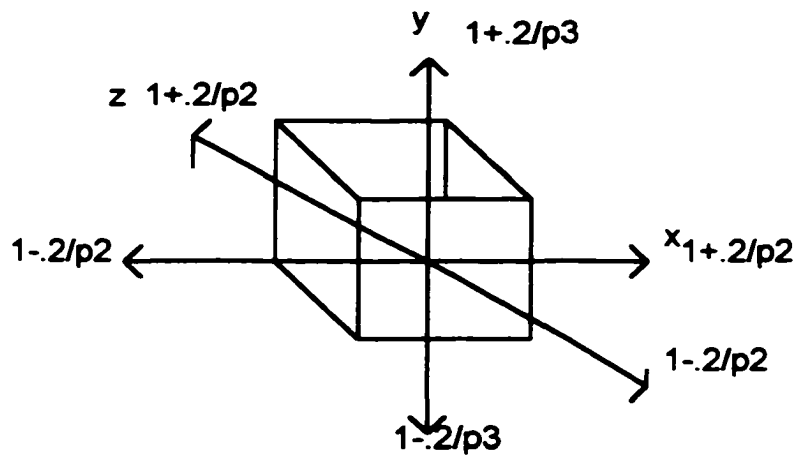
The 90% confidence regions using the sample data values from the data set with treatments, CAMF, CAP, and CA are as follows:

Ratio	Confidence Region
p_1/p_2	(1.554,1.785)
p_1/p_3	(4.009, 4.471)
p_3/p_2	(.271, .509)

Axes assumptions: x-axis p1 > p2

y-axis p1 > p3

z-axis p3 > p2



The endpoints of the 1-.2/pi side, where i= 2,3, of the axes are not defined under the assumptions. Hence, we can say that the left bounds of the posterior ratio violate the *equivalence cube* bounds. The pi's are estimated.

APPENDIX O

EQUIVALENCE CUBE FOR $\alpha = 3, \beta = 2$

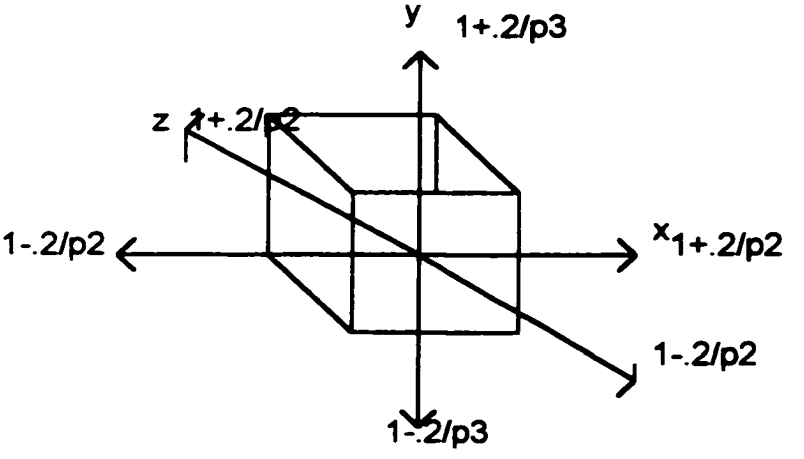
The 90% confidence regions using the sample data values from the data set with treatments, CAMF, CAP, and CA are as follows:

Ratio	Confidence Region
p_1/p_2	(1.555,1.785)
p_1/p_3	(4.007,4.473)
p_3/p_2	(0.270,0.510)

Axes assumptions: x-axis p1>p2

y-axis p1> p3

z-axis p3>p2



The endpoints of the 1-.2/pi side, where i= 2,3, of the axes are not defined under the assumptions. Hence, we can say that the left bounds of the posterior ratio violate the *equivalence cube* bounds. The pi's are estimated.

APPENDIX P

CONFIDENCE REGION FOR PAIRWISE COMPARISON OF THREE PROPORTIONS (USING BAYES' ESTIMATES AND SAMPLE DATA)

The following data are used in the tabulated results below:

	CR	No Response(%)	Total
PA	13	7	20
CT	12	8	20
ON	11	9	20
Total	36	24	60

The Bayes estimates are as follows:

α	β	$\frac{\hat{p}_1}{\hat{p}_2}$	$\frac{\hat{p}_1}{\hat{p}_3}$	$\frac{\hat{p}_3}{\hat{p}_2}$
2	3	1.07	1.15	.929
3	2	1.06	1.14	.930

where p_1 is the proportion of PA , p_2 is the proportion of CT and p_3 is the proportion of ON.

Table P1

Pairwise Comparison of Three Proportions (α = .1)
Confidence Regions for the Posterior of the Ratio of Proportions Using Bayes' Data

p_1/p_2	p_1/p_3	p_3/p_2	α	β
(0.889,1.205)	(0.972,1.328)	(0.712,1.146)	2	3
(0.888,1.232)	(0.987,1.293)	(0.727,1.133)	3	2

The following data are used in the tabulated results below:

	CR	No Response(%)	Total	p_i i=1,2,3
PA	13	7	20	0.65
CT	12	8	20	0.60
ON	11	9	20	0.55
Total	36	24	60	

The proportions from the sample data are as follows:

$\eta_1 = p_1 / p_2 = 1.08$

$\eta_2 = p_1 / p_3 = 1.18$

$p_3 / p_2 = .92$

where p_1 is the proportion of PA, p_2 is the proportion of CT and p_3 is the proportion of ON. The p's are estimated.

Table P2

Pairwise Comparison of Three Proportions ($\alpha = .1$)
Confidence Regions for the Posterior of the Ratio of Proportions Using SampleData

p_1 / p_2	p_1 / p_3	p_3 / p_2
(0.899,1.260)	(1.002,1.358)	(0.703,1.137)
(0.908,1.252)	(1.127,1.333)	(0.717,1.123)

Each region falls within the equivalence region. The confidence regions for the ratios of proportion p_1 / p_2 and p_3 / p_2 fall within the equivalence region, (0.667, 1.333). Also, the confidence region for the ratio p_1 / p_3 falls within the equivalence region, (0.636,1.364).

APPENDIX Q

EQUIVALENCE CUBE FOR $\alpha = 2, \beta = 3$

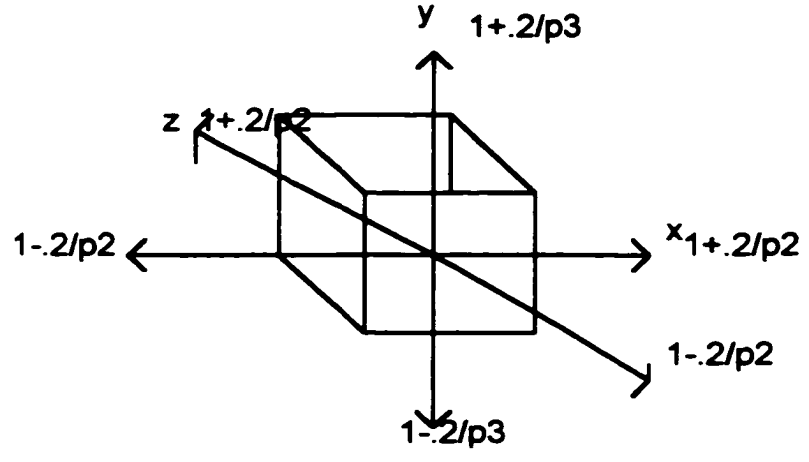
The 90% confidence regions using the sample data values from the data set with treatments, PA, CT, and ON are as follows:

Ratio	Confidence Region
p_1/p_2	(0.899,1.260)
p_1/p_3	(1.002, 1.358)
p_3/p_2	(0.703, 1.137)

Axes assumptions: x-axis p1>p2

y-axis p1> p3

z-axis p3>p2



The endpoints of the confidence region fall with the limits of the axes for each proportion. Hence, equivalence is established. The p's are estimated.

APPENDIX R

EQUIVALENCE CUBE FOR $\alpha = 3, \beta = 2$

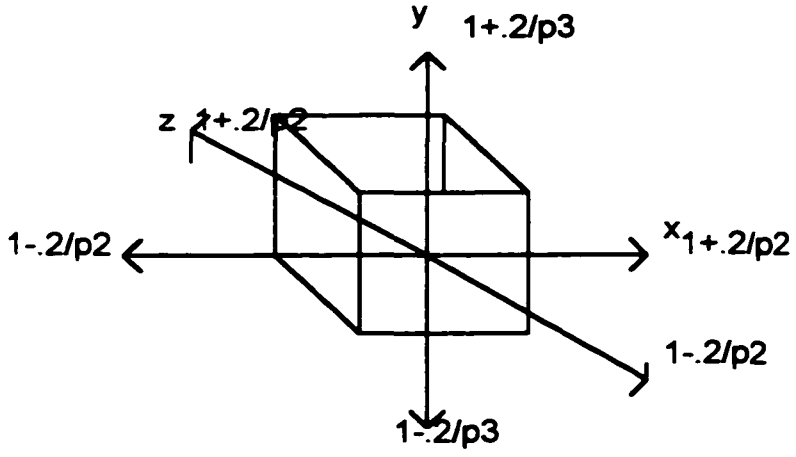
The 90% confidence regions using the sample data values from the data set with treatments, PA, CT, and ON are as follows:

Ratio	Confidence Region
p_1/p_2	(0.908,1.252)
p_1/p_3	(1.127, 1.333)
p_3/p_2	(0.717, 1.123)

Axes assumptions: x-axis p1>p2

y-axis p1> p3

z-axis p3>p2



The endpoints of the confidence region fall with the limits of the axes for each proportion. Hence, equivalence is established. The p's are estimated.

APPENDIX S

WEIGHTED LIKELIHOOD RATIO RESULTS FOR H_o: $\eta_1 = \eta_2 = 1$

Table S1 shows the results of the calculations of the weighted likelihood ratio (WLR) for several values of α and β. Remember that $\eta_1 = p_1/p_2$ and $\eta_2 = p_1/p_3$. The table was calculated using the data set consisting of treatments CAMF, CAP, and CA.

Table S1

Weighted Likelihood Ratio Results for $H_o: \eta_1 = \eta_2 = 1$

α	β	WLR
2	3	1.04×10^{-7}
3	2	4.60×10^{-1}

APPENDIX T

WEIGHTED LIKELIHOOD RATIO RESULTS FOR $H_o: \eta_3 = 1$

Table T1 shows the results of the calculations of the weighted likelihood ratio (WLR) for several values of α and β. Remember $\eta_3 = p_3/p_2$. The table was calculated using the data set consisting of treatments CAMF, CAP, and CA.

Table T1

Weighted Likelihood Ratio Results for $H_o: \eta_3 = 1$

α	β	WLR
2	3	5.60×10^{-6}
3	2	1.85×10^{-5}

APPENDIX U

WEIGHTED LIKELIHOOD RATIO RESULTS FOR $H_o: \eta_1 = \eta_3 = 1$

Table U1 shows the results of the calculations of the weighted likelihood ratio (WLR) for several values of α and β. Remember that $\eta_3 = p_3/p_2$ and $\eta_1 = p_1/p_2$. The table was calculated using the data set consisting of treatments CAMF, CAP, and CA.

Table U1

Weighted Likelihood Ratio Results for $H_o: \eta_1 = \eta_3 = 1$

α	β	WLR
2	3	3.10×10^{-3}
3	2	2.20×10^{-3}

APPENDIX V

WEIGHTED LIKELIHOOD RATIO RESULTS FOR $H_o: \eta_3 = \eta_2 = 1$

Table V1 shows the results of the calculations of the weighted likelihood ratio (WLR) for several values of α and β. Remember that $\eta_3 = p_3/p_2$ and $\eta_2 = p_1/p_3$. The table was calculated using the data set consisting of treatments CAMF, CAP, and CA.

Table V1

Weighted Likelihood Ratio Results for H_o: $\eta_2 = \eta_3 = 1$

α	β	WLR
2	3	7.48×10^{-10}
3	2	7.27×10^{-11}

APPENDIX W

WEIGHTED LIKELIHOOD RATIO RESULTS FOR COMPARISON

The following data are used in the following tabulated results below:

	CR	No Response(%)	Total
PA	13	7	20
CT	12	8	20
ON	11	9	20
Total	36	24	60

Remember that $\eta_3 = p_3/p_2$, $\eta_2 = p_1/p_3$ and $\eta_1 = p_1/p_2$.

Table W1

Weighted Likelihood Ratio Results for H_o: $\eta_2 = \eta_1 = 1$

α	β	WLR
2	3	98.72
3	2	35.12

Table W2

Weighted Likelihood Ratio Results for H_o: $\eta_3 = \eta_1 = 1$

α	β	WLR
2	3	63.46
3	2	13.15

Table W3

Weighted Likelihood Ratio Results for $H_o: \eta_2 = \eta_3 = 1$

α	β	WLR
2	3	141.23
3	2	123.20

Table W4

Weighted Likelihood Ratio Results for $H_o: \eta_3 = 1$

α	β	WLR
2	3	2.51
3	2	4.33

GRADUATE SCHOOL
UNIVERSITY OF ALABAMA AT BIRMINGHAM
DISSERTATION APPROVAL FORM
DOCTOR OF PHILOSOPHY

Name of Candidate Delicia Evet Carey

Major Subject Biostatistics

Title of Dissertation A Bayesian Approach for Assessment of Therapeutic Equivalence of Three Proportions

I certify that I have read this document and examined the student regarding its content. In my opinion, this dissertation conforms to acceptable standards of scholarly presentation and is adequate in scope and quality, and the attainments of this student are such that she may be recommended for the degree of Doctor of Philosophy.

Dissertation Committee:

Name		Signature
Alfred Bartolucci	**Co-Chair**	
Karan Singh	**Co-Chair**	
Deodutta Roy		
Heidi Weiss		
O. Dale Williams		

Director of Graduate Program

Dean, UAB Graduate School

Date 6/5/02