Statistical Analysis of Bioequivalence Studies

School of Statistics and

Actuarial Science

By

Mavuto Nyathi

576994

Supervisor

Renette Krommenhoek

A Research Report submitted to the Faculty of Science in partial fulfilment of the requirements for the degree of Master of Science

26 October 2016

DECLARATION

I declare that this Research Report is my own unaided work. It is being submitted for the Degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

ABSTRACT:

The cost of healthcare has become generally expensive the world over, of which the greater part of the money is spent buying drugs. In order to reduce the cost of drugs, drug manufacturers came up with the idea of manufacturing generic drugs, which cost less as compared to brand name drugs. The challenge which arose was how safe, effective and efficient the generic drugs are compared to the brand name drugs, if people were to buy them. As a consequence of this challenge, bioequivalence studies evolved, being statistical procedures for comparing whether the generic and brand name drugs are similar in treating patients for various diseases. This study was undertaken to show the existence of bioequivalence in drugs. Bioavailability is considered in generic drugs to ensure that it is more or less the same as that of the original drugs by using statistical tests. The United States of America's Food and Agricultural Department took a lead in the research on coming up with statistical methods for certifying generic drugs as bioequivalent to brand name drugs. Pharmacokinetic parameters are obtained from blood samples after dosing study subjects with generic and brand name drugs. The design for analysis in this research report will be a 2×2 crossover design. Average, population and individual bioequivalence is checked from pharmacokinetic parameters to ascertain as to whether drugs are bioequivalent or not. Statistical procedures used include confidence intervals, interval hypothesis tests using parametric as well as nonparametric statistical methods. On presenting results to conclude that drugs are bioequivalent or not, in addition to hypothesis tests and confidence intervals, which indicates whether there is a difference or not, effect sizes will also be reported. If ever there is a difference between generic and brand name drugs, effect sizes then quantify the magnitude of the difference.

KEY WORDS:

bioequivalence, bioavailability, generic (test) drugs, brand name (reference) drugs, average bioequivalence, population bioequivalence, individual

bioequivalence, pharmacokinetic parameters, therapeutic window, pharmaceutical equivalence, confidence intervals, hypothesis tests, effect sizes.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Mrs Renette Krommenhoek for the valuable assistance rendered to me throughout the writing of this project.

Thank you to the lecturers in the School of Statistics for teaching me various statistical techniques when analysing data, especially the aspect of interpreting computer outputs.

I would also like to thank my wife, Ofina Nyathi and children Sibusisiwe and Thabani for enduring long lonely periods of time when I was spending most of my time in the library compiling this report.

Last but not least, thank you Matthew Marowa (my nephew) for assisting and encouraging me to keep going even when I had family problems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOTATION AND TERMINOLOGY

$AUC_{0-\infty}$	Area Under a Curve from time zero to infinity.
AUC_{0-t}	Area Under a Curve from time zero to time t.
ABE	Average Bioequivalence.
ANDA	Abbreviated New Drug Application.
ANOVA	Analysis of Variance.
API	Active Pharmaceutical Ingredient.
AUC	Area Under a Curve.
BE	Bioequivalence.
C_{last}	The last amount of a Treatment Formulation.
C_{max}	The maximum quantity of a drug in the body.
C.I.	Confidence Interval.
CFR	Code of Federal Regulation.
DF	Degrees of Freedom in ANOVA.
ED	Estimates Distance Test.
ES	Effect Sizes.
FDA	Food and Drug Administration.
GDAC	Generic Drugs Advisory Committee.
HT	Hotelling T^2 Statistic.
IBE	Individual Bioequivalence.

ICC	Intraclass Correlation Coefficient.
IR	Immediate Release.
ln	Natural Logarithm.
LB	Lower Bound of the Confidence Interval.
LD	Likelihood Distance Test.
MCC	Medicines Control Council of South Africa.
MEL	Minimally Effective Level.
MM	Method of Moments.
MS	Mean-Shift Test.
MSE	Mean Square Error.
MTL	Maximally Tolerated Level.
OR	Odds Ratio.
PBE	Population Bioequivalence.
REML	Restricted Maximum Likelihood Method.
RR	Relative Risk.
RT	Reference followed by Test Treatment Formulation.
R^2	Coefficient of Determination.
SADC	Southern African Development Community.
SE	Standard Error.

$T_{\frac{1}{2}}$	The amount of time taken by a drug in the body to decrease by a half.
T_{max}	Time from administering the drug, to reach C_{max}.
TOST	Two One-Sided Tests.
TR	Test then Reference Treatment Formulation.
UB	Upper Bound of the Confidence Interval.
US	United States.
USA	United States of America.
V	Volume of Distribution.
λ	Is a constant where the drug is excreted from the body over a period of time.

Chapter 1

Introduction

1.1 General Introduction

The cost of healthcare is high mainly due to expensive drugs. Generic drugs (test treatment formulation) are manufactured, with some stages which are carried out when manufacturing brand name drugs (reference treatment formulations) excluded, for example, expensive clinical trials. As a result, generic drugs cost less as compared to the brand name drugs hence reducing healthcare costs. During drug development bioequivalence studies are used. This chapter motivates the need for carrying out bioequivalence studies. In bioequivalence studies the test treatment and reference treatment formulations are compared to check if they are bioequivalent.

Definitions of generic and brand name drugs, bioavailability and bioequivalence are discussed in this chapter. Relevant methods for bioequivalence studies, pharmacokinetics and pharmacodynamics are mentioned. The source of data and data collection method is stated. The outline of the rest of this research report completes this chapter.

1.2 Background Information

Bioequivalence studies are used in the industries where drugs are manufactured, pharmaceutical industries, during the development of new and generic drugs. The basis of this bioequivalence study was to compare a drug product to be tested (generic drug) with an appropriate reference treatment formulation (brand name or original drug or innovator drug). The cost of healthcare is expensive, with the main contributing factor to the high cost of healthcare being the cost of drugs according to many researchers who include, Chow and Liu (2000), Meredith (2003), Midha and Mackay (2009), Kamerow (2011), Mastan, Latha and Ajay (2011) and Qayyum (2012). Generic drugs are manufactured to try and reduce the cost of drugs in the market. They are cheaper than the brand name drugs but the quality should be almost the same as that of the reference drugs, in terms of the

safety and effectiveness in treating ailments. Maintaining the quality of generic drugs is mentioned in many research papers, though, consumers are concerned about that generic drugs may be bioequivalent but not necessarily therapeutic equivalent. To avoid this fear many guidelines or guidance and regulations on bioequivalence covering the licensing of generic products have been published as mentioned in Section 2.2 on literature review to ensure that the market dispenses quality drugs.

Therapeutic equivalence refers to the effectiveness and safety of the generic drugs in producing similar results as compared to the reference drugs in terms of treating diseases. The Medicines Control Council of South Africa (MCC, 2003) guideline to bioequivalence defines therapeutic equivalence as two pharmaceutical products that are therapeutically equivalent after same quantities of a drug have been administered to patients, regarding both their efficacy and safety, as determined from an appropriate bioequivalence, pharmacodynamics, clinical or in vitro studies. Southern African Development Cooperation (SADC, 2007) guidelines to bioequivalence defines therapeutic equivalence in a similar manner compared to the MCC (2003) since they both talk of that two pharmaceutical drugs, the generic and reference should be more or less equally safe and effective, after administration of a drug in the same molar dose, as determined from an appropriate bioequivalence method.

1.2.1 Test and Reference Treatment Formulations

A test treatment formulation is a drug which has more or less the same chemical composition and is as safe and effective, after it has been administered to a patient, as the reference drug. A reference treatment formulation (brand name drug) is the original drug formulation from which the test (generic) is developed. Meredith (2003) stated that a test treatment drug must contain similar amounts of the same active ingredient in the same formulation and route of administration as compared to the reference treatment formulation. Dighe (1999) also looks at generic drug formulations as drugs which when prescribed to patients are the same as brand name drugs in terms of possessing the identical active drug

substance in the same amount or concentration and the same dosage form after being administered by the same route of administration as the brand name drug products. Mastan *et al.* (2011) contrasted the terms brand name and generic drug, by referring to a brand name drug as a formulation available in the market, sold under a known trademark protected name while a generic is the same as a brand name in dosage, safety, strength and its intended use. According to Qayyum (2012) the reference treatment formulation is the compound that was developed first and was approved to be marketed for purposes of treating health conditions after it showed satisfactory efficacy and safety. The generic drug is a result of the modification which is done on the reference drug though preserving the quality such that both drug formulations are pharmaceutically equivalent and the same in terms of how they are taken, as well as quality and performance. SADC (2007) defines a reference treatment formulation as a pharmaceutical control drug product to be compared to a new product being developed. Both drug formulations should produce similar effects after being administered, in equal quantities and using the same route of administration, in terms of efficacy, safety and quality. Based on the above definitions of generic and brand name drugs by Meredith (2003), Dighe (1999) and Mastan *et al.* (2001), it is evident that in bioequivalence studies generic drugs are tested to check whether they are safe and effective for treating diseases as compared to the brand name drugs. In South Africa the Medicines Control Council is a body which is responsible for ascertaining drug formulations as bioequivalent or not. MCC (2003), on discussing generic and reference drugs, indicated that the reference product must be a drug available in South African pharmacies. The generic drug is then manufactured modelling the reference product on the South African market so as to ensure that both drugs are pharmaceutically equivalent, hence quality, effectiveness and safety is maintained.

1.2.2 Bioavailability

A bioequivalence study allows researchers to compare bioavailability between the generic drug and the reference drug to find out if there is a significant difference between the two formulations. Qayyum (2012) define bioavailability as the

concentration of a drug that is in the blood after dosing. Bioavailability according to the definitions provided by the MCC (2003) and SADC (2007) refers to the rate and extent to which the active pharmaceutical ingredient (API) is absorbed from a pharmaceutical product and becomes available at the appropriate part of the body. This definition of bioavailability correlates with one provided by the Food and Drug Administration (FDA, 2003) which states that bioavailability is the rate and extent to which the active ingredient of a drug formulation is absorbed from the drug product and becomes available at the site of action. In general, bioavailability is whereby a drug formulation administered is absorbed in the body and become available where it is intended to be used.

1.2.3 Bioequivalence

A generic drug is said to be bioequivalent if its difference in terms of bioavailability is minimal as compared to a reference drug when evaluated in similar conditions. SADC (2007) defines bioequivalence as an insignificant difference in terms of bioavailability between two pharmaceutically equivalent products or pharmaceutical alternatives under similar conditions in an appropriately designed study. FDA (2003) defines bioequivalence as a small difference in the rate and extent to which the active ingredient in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered using same quantities of the drug under similar conditions in an appropriately designed study. Rani and Pargal (2004) and Lopes (2009) also define bioequivalence as an insignificant difference between the brand name and generic drug formulations in terms of their bioavailability. The basis of bioequivalence is that when two treatment formulations of the same drug, generic and reference, are equivalent in the rate and extent of their drug absorption, they should have the same therapeutic effect. In practice, we can never have two treatment formulations with exactly the same bioavailability, an insignificant difference would always exist.

1.2.4 Pharmacokinetic and Pharmacodynamics Studies

Pharmacokinetic and pharmacodynamics studies are used in conjunction to establish bioequivalence. Pharmacokinetics is the study of the way in which drugs move through the body, that is, what the body does to the drug while pharmacodynamics account for the effect that drug(s) have on the body. Pharmacokinetics is a science describing drug:

- Absorption from the administration site;
- Distribution to, tissues and target sites of desired and /or undesired activity;
- Metabolism;
- Elimination or excretion (Peer, 2007).

1.3 Significance of the Research

Brand name drugs and generic drugs are both available in the market, but the brand name drugs are usually expensive compared to the generic. The cost of healthcare has been rising in the last two decades and the main aspects causing the increase according to Chow and Liu (2000), Borgherini (2003), Meredith (2003), Midha and Mckay (2009), Kamerow (2011), Mastan *et al.* (2011) and Kalpesh, Sokindra and Kishore (2013) are the expensive drugs. The cost of drugs, especially the original ones, makes medical expenses escalate to a point where the majority of people cannot afford it. Hence, high costs of drugs necessitate bioequivalence studies, as there is a need to find drugs which cost less as compared to the brand name ones. Generic drugs cost less than the original drugs though their effect on treating patients is similar to that of the original drugs. Generic drugs are developed at a lower cost because some stages or tests which were done when manufacturing brand name drugs are skipped now, for example, expensive chemical trials are not necessary. Kamerow (2011) states that generic drugs which can be procured at a lesser cost as compared to brand name drugs are available in the market for a wide range of major diseases including diabetes, hypertension, heart failure, lipid disorders and acid reflux. Bioequivalence studies are an area of research pursued by many researchers in recent years because there

is a need to reduce the financial burden of high medical costs due to expensive drugs. In addition to being bioequivalent, the pharmaceutical effectiveness of a generic drug should also be of almost the same degree with the reference drug so as to preserve the quality of a generic drug. The stages carried out in bioequivalence studies involve: designing the study, conducting the study and the evaluation of results.

It has been observed by Kamerow (2011) that in the United States of America (USA), healthcare consumes a big part of the country budget, approximately 10% of the budget. He goes on to indicate that, it is observed that changing from using original drugs to generic drugs seems to be a feasible way of reducing the costs of the healthcare budget and quotes figures to substantiate the claim that costs are really reduced by switching to generic drugs. The USA government was able to save $33bn in 2007 by adopting a programme called Medicare, where generic drugs are prescribed to patients. Overall savings from the healthcare budget were estimated to be $139 in 2009 from the allocated amount by using the generic prescriptions in the United States (US). It is further indicated in the article by Kamerow (2011) that the amount saved in 2009 was not much significant contrasted with the $2.5 trillion annual healthcare bill, but, it should be noted that such reductions will eventually go a long way in saving some funds on healthcare. Mastan *et al.* (2011) had the same observation and opinion as Kamerow (2011) on that a good strategy for lowering healthcare costs is to introduce generic drugs which are much cheaper compared to the brand name drugs. This strategy has been effective in reducing total prescription cost by 11% without sacrificing quality. Generic drugs have captured more than 65% of the global market and account for 66% of prescriptions filled in the US but for less than 13% of the cost. Borgherini (2003) is also of the opinion that healthcare costs can be drastically reduced by introducing generic drugs in the market. Many countries including those in the industrialized world have a challenge of incurring high healthcare costs, hence they realise that there is a need to substitute the brand name drugs by generic drugs. He indicates that drug patents in the European Union expire after six to ten years after registration, after which generic drugs are then introduced in

the market at a lesser cost because registration costs of original drugs are not carried over to the generic ones.

Advantages of using generic drugs are:

- There is more knowledge about the chemical composition as well as the therapeutic effects of a generic drug as compared to a brand name drug;
- A uniform name can be chosen for a test treatment formulation when it gets distributed into the market;
- Pharmacists have the freedom to select the most suitable formulation of generic drug in terms of quality and price (Meredith, 2003).

However, Meredith (2003) just like many researchers on bioequivalence mentions that the safety and effectiveness of generic drugs need to be monitored. The disadvantage highlighted is that the rate and extent of absorption may differ between test and reference treatment drugs and that the physical properties of the generic drug such as, colour, shape, size and flavour might not be the same as that of the brand name drug. Borgherini (2003) and Mastan *et al.* (2011) also note that the reduction in costs of drugs due to the use of generic ones is undisputed, though there are disadvantages of using generic drugs, they discuss three disadvantages. The first disadvantage is that the difference in costs of purchasing generic drugs is not much between those using medical aid compared to those who are not. There is therefore the non-cost benefit felt by those using medical aid. Secondly, the clinical equivalence of generic drugs as compared to brand name drugs is not fully investigated. FDA (2001) approves the sale of generic drugs which are bioequivalent to be within 90% confidence interval, this reduces costs but as argued by many researchers, clinical equivalence is not fully assessed. Thirdly, confusion arises, especially with elderly patients when packaging changes, for example, if a pill which used to be yellow and round is now dispensed as a green pill in the form of an oblong shape after a generic drug is introduced or changed. According to the observations by Kamerow (2011) and Midha and McKay (2009) it is quite evident that the use of generic drugs have pros and cons, but the fact that they are lower in costs and their drug safety and effectiveness is similar to the original drugs necessitates further research to be done in addressing the

disadvantages. In addition to being bioequivalent, the pharmaceutical effectiveness of a generic drug should also be of almost the same degree with the reference drug so as to preserve the quality of a generic drug.

1.4 Aims and Objectives

1.4.1 Aims

The aims of the research were to:

- Establish whether the test treatment formulations are safe and efficient as compared to the reference drugs when administered to patients;
- Ensure that the test treatment formulation are a copy of the reference treatment drugs with the same dosage, strength, route of administration as well as its intended use;
- Justify why bioequivalence studies are necessary;
- Provide areas for further research.

1.4.2 Objectives

The objectives of this study were to assess bioequivalence between a generic drug and a brand name drug (an antibiotic) using statistical tests. Pharmacokinetics, pharmacodynamics, clinical studies and in-vitro as well as in-vivo studies are suggested by Mastan *et al.* (2011) for the purpose of assessing bioequivalence.

1.5 Research Instruments

Blood samples obtained through a 2×2 crossover design from fourteen sheep were used to get the data for the research project. The data set was obtained courtesy of the University of Pretoria at Onderstepoort within the Directorate of Veterinary Pharmacology. It consists of fourteen sheep, each being dosed by a test and reference treatment formulation in phase one (sequence one) then reference and test treatments in phase two (sequence two). The drugs were administered intra muscular, whereby the original drug was an antibiotic. The antibiotic was used to test the effect of heart water (parasite infection) and 10 mg/kg was administered per time interval.

The drugs were tested on sheep to evaluate bioequivalence hoping that if they are bioequivalent chances are likely that if administered on human beings the same effect would be noted. When they are tested on human beings results would have already been shown that the drugs are bioequivalent, such that the main focus would be to check if the same effect of the drugs is achieved.

The software to be used will be a combination of Microsoft Excel (2010), SAS Enterprise Guide 7.1 and SAS 9.4.

1.6 Outline of the Rest of the Project

The next section of the research, Chapter 2, looks at the literature review. In this chapter, the evolution of bioequivalence studies is traced and then the fundamental concepts are defined and discussed, notably:
Average bioequivalence (ABE) assessed by the confidence interval approach and Schuirmann's interval hypothesis testing using two one-sided tests procedures on the pharmacokinetic parameters, area under a curve (*AUC*) and C_{max} for the test and the reference treatment drugs. Wilcoxon Mann-Whitney two one-sided test is used to analyse the bioequivalence for the parameter T_{max}. Population bioequivalence (PBE) and individual bioequivalence (IBE) are also discussed. The procedure recommended by the Food and Drug Administration (FDA, 2001) is used to assess PBE. Chapter 3 describes the methodology, which includes the study design to be used in this research report, 2×2 crossover design, pharmacokinetic parameters needed for the bioequivalence study. Parametric, nonparametric distributions are discussed. Outliers, power of a test, bootstrapping and effect sizes are also covered. Chapter 4 deals with the results and analysis of results achieved after using statistical models/techniques to establish bioequivalence between test and reference drugs. Chapter 5 discusses the conclusion and recommendations based on the research.

Chapter 2

Literature Review

2.1 Introduction

In this chapter, a review of bioequivalence studies from where they started until current time is undertaken. Pharmacokinetic parameters, notably, AUC_{0-t}, $AUC_{0-\infty}$, $C_{max}$and T_{max} are the metrics which are used to compare bioavailability between the test and reference treatment formulations. AUC_{0-t}, $AUC_{0-\infty}$ and C_{max} should be logarithmically transformed before statistical data analysis is carried out, while, T_{max} is analysed on the original scale of measurement. The justification for transforming the data is discussed in this chapter as well as the criteria used to certify drugs as bioequivalent. This chapter also looks at bioequivalence models and methods of assessing ABE, PBE and IBE.

2.2 A Short Account of the Beginnings of Bioequivalence Studies

The Americans were pioneers of this area of drug development leading to marketing the final product. The Food and Drug Administration, the research centre responsible for checking and ascertaining that the test treatment and reference treatment formulations are bioequivalent took a considerably amount of time, dating back from the early 1970s to come up with statistical models of ABE, PBE and IBE which are generally accepted worldwide. Other countries also set up their drug development centres emulating the Americans (Jones and Kenward, 2003). Midha and Mckay (2009) also indicated that bioequivalence studies have been ongoing for over twenty years and are accepted by pharmaceutical industries as a method of approving generic drug products which are sold at a reduced cost compared to the brand name drugs.

Chow and Liu (2000) trace the evolution of bioequivalence studies from its early days and state that the research and developments on bioequivalence studies commenced from the early 1970s and are in four stages. The first stage was from the early 1970s to 1984, when the Drug Price Competition and Patent Restoration Act (1984), was passed. The next stage was from 1984 to 1992, when a

bioequivalence guideline was published: FDA (1992) Guidance on Statistical Procedures for Bioequivalence Using a Standard Two-Treatment Crossover Design. The third phase, where population and individual bioequivalence and their statistical procedures were discussed kicked off from 1992. The fourth phase started from the beginning of the twenty first century, basically reviewing and updating the research conducted in the last thirty years of the twentieth century. Another guideline on BE studies, FDA (2001) Guidance for industry: Statistical Approaches to Bioequivalence, was issued out during this period.

During the early 1980s focus shifted from bioavailability, to bioequivalence though the two studies still complement each other and are assessed through statistical methods. The Drug Price Competition and Patent Term Restoration Act of 1984 whereby generic drug products could only be approved through bioavailability and bioequivalence studies was a noble one because there is need to find the balance between producing drugs which cost less compared to the original ones while the quality of the drugs need not be compromised to ensure safety and effectiveness when administered to patients. In 1986 concerns were raised over the safety and efficacy of generic drugs. The challenge was that generic drugs could match the required standard for them to be approved as bioequivalent but therapeutic equivalence was not fully investigated. There was a need, through statistical tests or evidence to assure people that the bioequivalent drug products were also therapeutically equivalent. Midha and Mckay (2009) pointed out that issues or challenges on whether bioequivalent drug products are also therapeutic equivalent were addressed by the FDA (2003) during a meeting on bioequivalence of solid oral dosage forms.

Since certain acceptable standards have to be attained in order to approve a generic drug, it happened that when bioequivalence studies started to gain momentum, regulatory authorities were also put in place to approve or disapprove the generic products. In US, the FDA published guidelines on how industry can properly use bioavailability and bioequivalence studies well, which were entitled Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drugs General Considerations. Such guidelines are continuously

being improved or updated to meet the current changes in pharmaceutical needs and technology. Other FDA guidelines were in 1992, 1999, 2001 and 2003 and then elsewhere following the research done in US, regulatory authorities were established which published guidelines to be followed if a generic drug is to be licensed such that it becomes available on the market at a lower cost as compared to the brand name drug.

In South Africa generic drugs are controlled by the medicines regulatory authority, known as the Medicines Control Council (MCC, 2003), which was established in the early 1970s. The MCC (2003) is a body that was constituted under the Medicines and Related Substances Control Act, 101 of 1965, to monitor the regulation of medicines in South Africa. The members of this statutory body are chosen by the Minister of Health and its major function is to safeguard and protect the public through ensuring that all medicines that are marketed and used in South Africa are safe, therapeutically effective and are consistently of the required standard of quality.

Regulatory authorities established in other countries include Japanese guidelines (2001), MCC (2003), Central Drugs Standard Control Organization (CDSCO, 2005) and SADC (2007). These regulatory authorities concur on that generic products have to be pharmaceutically equivalent and should be of the same bioavailability as compared to the original drugs after administration. There is also a criterion, pharmacokinetic parameters, which they use to certify a test treatment formulation as bioequivalent to a reference treatment formulation.

2.3 Pharmacokinetic Parameters

Wang and Bakhai (2006) recommended that statistical analysis of bioequivalence should be based on non-compartmental pharmacokinetic parameters, namely:
AUC: refers to the area under a curve, it depicts a blood concentration time curve which illustrates information on the extent of the absorption of a drug, be it generic or brand name. Appendix 1 has figures showing *AUC* for the sheep used in this research report.

AUC_{0-t}: is the area under the plasma/serum/blood concentration time curve from time zero to time *t,* where *t* is the last time point where the amount of drug concentration can be measured.

$AUC_{0-\infty}$: is the area under the blood concentration time profile from time zero to infinity.

C_{max}: is the maximum value of drug concentration.

T_{max}: is the time point corresponding to C_{max}.

λ is the elimination constant which describes the loss of drug activity from the body per time unit (for example, per hour). Chow and Liu (2000) states that λ is the constant rate at which the drug is excreted and can be estimated as the slope of the terminal part of the logarithmically transformed concentration time curve multiplied by –2.303. The justification of multiplying 2.303 by a minus is due to the fact that at the stage where the drug is excreted from the body (elimination phase), the slope of the line is negative.

$T_{\frac{1}{2}}$: refers to the time taken for the drug concentration to decrease by half when a drug is eliminated from the body.

Regulatory guidelines on bioequivalence studies recommends that all concentration dependent pharmacokinetic parameters, for example, *AUC* from time zero to time *t, AUC* from time zero to infinity and C_{max} should be logarithmically transformed using either common logarithms to base 10 or natural logarithms (base *e*). The choice of adopting the use of common or natural logarithms should be consistent and should be stated in the bioequivalence study report (FDA, 1992). The view of transforming the bioequivalence pharmacokinetic parameters *AUC* and C_{max} is also expressed by Concordet (2004) stating that transforming the parameters prior to the analysis enables the researcher to proceed using assumptions of normality. He also indicates that the logarithmic transformation of the parameters C_{max} and *AUC* helps to make the variance constant so as to obtain a symmetric distribution but the parameter T_{max} rarely follows a symmetric distribution and the variance remains unstable.

There are several reasons given motivating why the parameters *AUC* and C_{max} should be logarithmic transformed:

2.3.1 Clinical Rationale

In the USA a committee responsible for giving direction as to what needs to be done in bioequivalence studies, known as the Generic Drugs Advisory Committee (GDAC) agreed in a meeting which was held in September 1991 that the ratio of the means of two treatment formulations, test and reference, need to be considered for the evaluation of bioequivalence instead of the difference of the means. The ratio of means can only be compared statistically after the data has been logarithmic transformed (FDA, 1992).

2.3.2 Pharmacokinetic Rationale

The 2×2 crossover design used in bioequivalence studies consists of subjects, period, sequence and treatment effects. As a result, it is assumed to be a model where all these components are added to each other. However, since GDAC (1991) advised that the ratio of two treatment formulations be used to assess bioequivalence, the 2× 2 crossover design would then become a model where the various components, (subjects, period, treatment effects and sequence) are multiplied because of the logarithmic transformation to get the ratio of means (Rani and Pargal, 2004).

2.3.3 Statistical Rationale

Statistical analysis is done effectively and efficiently if the bioequivalence dataset approximately follows a log-normal distribution. In practice, the pharmacokinetic parameters *AUC* and C_{max} are usually skewed and there is a correlation between the means and the variances, whereby as the means of the treatment formulations increase, the variances also increase. Logarithmic transformation is carried out to reduce skewness in the parameters so that they become almost symmetrical and ensures that the increase of the means does not affect the variance (Concordet, 2004).

T_{max} is obtained from a set of discrete values, though not exact, they are measured with errors. Parametric statistical methods used in bioequivalence studies cannot be applied for the parameter T_{max}, as they are considered inappropriate (Hauck and Anderson, 1984). Nonparametric tests such as the Wilcoxon Signed-Ranks test, Wilcoxon Rank Sum test or Wilcoxon Mann-Whitney two one-sided tests seem appropriate for the analysis of T_{max} (Westlake, 1976).

2.4 Bioequivalence Criteria

The regulatory authorities indicate that generic products and brand name products are bioequivalent if the ratio of the means of C_{max} and *AUC* of the two treatment formulations should lie in the range 0.80 to 1.25. If the ratio is less than 0.80 we say there is sub-availability while greater than 1.25 implies super-availability. Schall and Endrenyi (2010) point out that the bioequivalence range need to be tightened from 0.90 to 1.111 especially for cases of formulations with narrow therapeutic index (window). T_{max} (time to reach maximum blood concentration) should also be similar between the generic product and the original product. Regulatory authorities FDA (2001) strive to ensure that the scientific or statistical approaches used for assessing bioequivalence are of an acceptable standard, as mentioned earlier, so that the quality of generic drugs, though at reduced costs, is high.

All stakeholders, at all possible levels, both nationally and internationally, have been convened in meetings, conferences and workshops to discuss as well as to reach agreements on approaches to standardise bioequivalence of pharmaceutical products in order to satisfy pharmaceutical equivalence. According to SADC (2007) the key issue is to ensure safety and efficacy of generic or new drugs in line with the FDA (2001) guidelines and regulatory bodies in other countries. Regulatory bodies concur that generic drugs could be bioequivalent but they should be also pharmaceutically equivalent at the site of action for a desired period of time.

2.5 Pharmacokinetic and Pharmacodynamics Models

Pharmacokinetics and pharmacodynamics techniques are used in the development of drugs to ascertain bioequivalence of new or generic drugs. Greenblatt, Moltke, Harmatz and Shade (1998) discuss pharmacokinetics and pharmacodynamics procedures noting that they have been applied during the process of the development of new drug entities as well as for the improved understanding of the clinical actions of drugs that are already marketed. Pharmacokinetics is based on mathematical equations that allow the prediction of a drug's behaviour and more emphasis is put on the relationships between drug concentrations and the time elapsed after administering the treatment formulation. According to the FDA (1999), population pharmacokinetics is the study which looks at the amount by which treatment formulation effects measured in a study subject differ after dosing with the test and reference treatment formulations., whereas pharmacodynamics measures the effect of drug concentration in individuals. In general, pharmacokinetics is the study of the way in which drugs move through the body during absorption, distribution metabolism and excretion and pharmacodynamics refers to the relationship between drug concentration at the site of action and the resulting effect including the time course and intensity of therapeutic and adverse effects. In short, pharmacokinetics refers to the effect the body has on the drugs whereas pharmacodynamics accounts for the effect that drugs have on the body. Models for pharmacokinetic and pharmacodynamics are used in the development of drugs and they complement each other in the sense that, pharmacokinetic models project the time period by which the treatment drug would be still available in the body after a treatment drug has been administered but the pharmacodynamics models accounts for the intensity of drug effects on subjects used for research, whether human or animals (Greenblatt *et al*, 1998). The combination of pharmacokinetics and pharmacodynamics give rise to kinetic-dynamic modelling which investigates the drug concentration-effect relationship. This relationship is important in bioequivalence studies because clinical therapeutic of a drug is determined by a concentration-effect relationship.

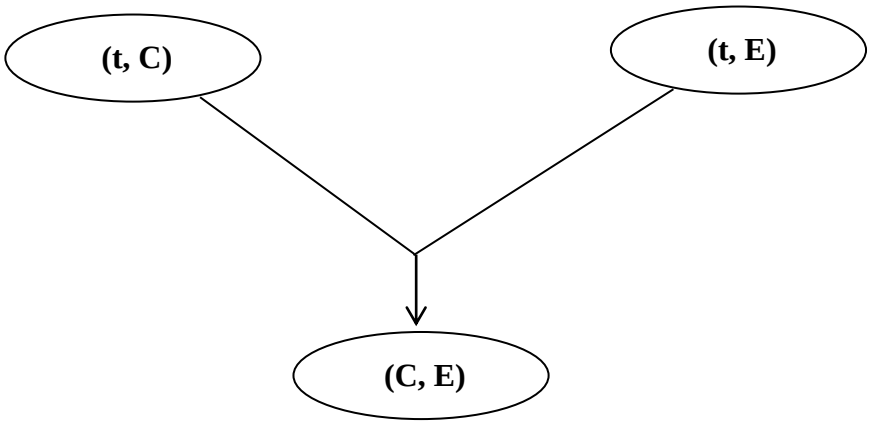


Figure 2.1 Concentration-Effect Model

The diagram above is adapted from Greenblatt *et al.* (1998)

Greenblatt *et al.* (1998) point out that some models have been developed which incorporates the concentration-effect relationship. The models are:

Sigmoid E_{max}: $E = \frac{E_{max} \cdot C^A}{C^A + EC_{50}^A}$ (2.5.1)

Exponential: $E = m.C^A$ (2.5.2)

Linear: $E = m.C$ (2.5.3)

where;

E: Effect,

C: Concentration,

E_{max}: is the maximum pharmacodynamics effect,

EC_{50}: is the 50% effective concentration,

C^A: the exponent A represents the steepness of the concentration-response relationship in ascending order, though the biological importance of A is not established.

The sigmoid E_{max} is an important model since conclusions can be drawn from it about the potency and efficacy of drugs producing the same clinical effect, individual differences in drug sensitivity, the mechanism of action of pharmacologic potentiators or antagonists, and the possible clinical role of new medications, these realisations were made by Greenblatt *et al.* (1998). The model has limitations, it does not apply to all concentration-effect data and an example of where the sigmoid E_{max} cannot be used is when the experimental data is not consistent with the model. Many researchers use both techniques, pharmacokinetic and pharmacodynamics in the development of generic drugs, of which nowadays the emphasis is more on the accuracy and sensitivity of the results when measuring drug concentration and drug effects respectively.

2.6 Compartmental and Non-Compartmental Models

Models have been developed that explain the interaction between an organism and drugs. There are two classes of models, namely: the non-compartmental and the compartmental models. According to Schutz (2009), compartmental models use kinetic models to describe and predict the amount of treatment drug in the body. Compartmental models are subdivided into non-compartmental, two compartment models and multi-compartment models. The non-compartment model (single compartment) considers a study subject (organism) as one homogeneous compartment whereby if a dose is administered, the blood concentration of a drug is distributed uniformly to other body fluids or tissues resulting in the excretion of the drug being directly proportional to the drug concentration in a subject under study. Peer (2007) describes the human body as a single compartment through which the drugs circulates through various organs of the body. A drug taken orally in tablet form is absorbed into this compartment as the tablet dissolves in the stomach and drug elimination from this compartment is due to the actions of the liver as well as kidneys. While the tablet is being dissolved in the stomach, the rate of absorption of the drug into the circulating blood is greater than the rate it is eliminated and the concentration of drug in the blood increases. After the tablet has been dissolved the rate of elimination is greater and the concentration declines.

A non-compartmental (single compartment) model measures the concentration of a drug administered by estimating the *AUC*. A numerical integration method, the trapezoidal rule is used to estimate the area under a curve of a concentration. The equation encompassing method of residuals is used to estimate the constant rate of absorption of a treatment formulation (Chow and Liu, 2000).

$$C_t = \frac{k_a\ FD_0}{V(k_a - k_e)}(e^{-k_e t} - e^{-k_a t}) \quad (2.6.1)$$

where;

k_a and k_e are the absorption and the constant rate of elimination, respectively,

D_o is the amount of drug administered,

V is the volume of distribution,

F is the fraction of the drug that penetrates the body up to where circulation takes place,

t is the time period for drug concentration in the body.

In equation 2.6.1 C_{max} and T_{max} can be similarly obtained as follows:

$$T_{max} = \frac{2.303}{k_a - k_e}\log(\frac{k_a}{k_e}) \text{ and} \quad (2.6.2)$$

$$C_{max} = \frac{k_a\ FD_0}{V(k_a - k_e)}(\mathrm{e}^{\mathrm{k_e tmax}} - \mathrm{e}^{\mathrm{k_a tmax}}). \quad (2.6.3)$$

In practice, C_{max} and T_{max} are not found by applying the formulae in equation 2.6.2 and 2.6.3 but are read from the blood concentration profiles.

$T_{\frac{1}{2}}$ is the amount of time taken for the drug concentration to decrease by a half is considered when a drug is eliminated from the body, computed as:

$$logD = logD_0 = \frac{-k_e t}{2.303} \quad (2.6.4)$$

where *D* is the quantity of drug in the body. Thus, at $D=\frac{D_0}{2}$, that is, $t=T_{\frac{1}{2}}$ we have

$$log(\frac{1}{2}) = \frac{-k_e T_{\frac{1}{2}}}{2.303} \quad (2.6.5)$$

Hence

$$T_{\frac{1}{2}} = \frac{0.693}{k_e} \quad (2.6.6)$$

where k_e is given by

$$k_e = (-2.303)(\frac{dLogD}{dt}) \quad (2.6.7)$$

The concentration-time graph of a one-compartment model is a linear graph or approximately linear. However, in reality, this model seems inadequate in pharmacokinetics since an organism's supply of blood is not the same for all parts, there are areas or parts of an organism which receive more blood supply as compared to others. A two compartment model comprises the central compartment, there is more blood supply, for example, liver and kidneys and in the peripheral compartment organs of the body have a lower blood flow, for example, the brain tissues. The elimination of a drug tends to be rapid for a two compartment model in the central compartment though in rare occasions it occurs in the peripheral compartment or even in both. The two compartment model, according to many researchers performs better than the non-compartmental model though in real life each body tissue has its own distribution characteristics, hence a multi-compartment model is suggested, which is represented by fitting a curve. Pharmacokinetics parameters, such as the area under the curve are then calculated. The compartmental models are able to predict the concentration of a treatment formulation at any time as compared to non-compartmental models but are difficult to develop and validate. Non-compartmental models do not have the limitation of having different compartments within an organism. Both are used in finding pharmacokinetics parameters (Peer, 2007).

2.7 Methods of Assessing Bioequivalence

2.7.1 Average Bioequivalence

Patterson (2010) points out that the FDA (1999) initially used average bioequivalence to assess the bioequivalence between brand name and generic drugs, though later on, there were debates as to whether certifying drugs using the average bioequivalence criterion was enough or not. It has been noted that average bioequivalence has limitations since it only compares the population means between test and reference treatment formulations. The intra subject variance of the formulations or the subject by formulation interaction is not taken into account. Hence, there is a major concern to know whether approved test treatment products can be used safely and interchangeably, if certified bioequivalent using

average bioequivalence. As a result of the limitations of average bioequivalence, the FDA (2001) recommended a statistical test procedure for PBE and IBE. Currently, researchers on bioequivalence of generic and original drugs combine population bioequivalence and individual bioequivalence for certifying formulations as bioequivalent or not.

The Average Bioequivalence criterion (ABE) is given as:

$(\mu_T - \mu_R)^2 \leq \theta_A^2$ (2.7.1.1)

also given as:

$-\theta_A \leq \mu_T - \mu_R \leq \theta_A$ (2.7.1.2)

where;

θ_A =The average bioequivalence limit which is $\theta_A = ln(1.25)$,

μ_T =Mean of the test treatment formulation,

μ_R =Mean of the reference treatment formulation,

$(\mu_T - \mu_R)$ =Mean difference of the two treatment formulations.

2.7.2 Confidence Interval Approach

Bioequivalence study data is best evaluated using 90% confidence intervals of the individual parameters of interest, usually, C_{max} and *AUC*, after transforming original data using logarithmic transformation based on common logarithms (base 10) or natural logarithms (base *e*). According to Chow and Liu (2000) the confidence interval approach is one of the many methods which can be used for certifying drugs as bioequivalent or not.

To calculate confidence intervals for untransformed data, we let;

$\bar{X}_{T1}$ =Mean of the test treatment formulation in period 1,

$\bar{X}_{T2}$ =Mean of the test treatment formulation in period 2,

similarly;

$\bar{X}_{R1}$ and $\bar{X}_{R2}$ =Mean of the reference treatment formulation in periods 1 and 2 respectively,

$\bar{X}_T = \frac{1}{2}(\bar{X}_{T1} + \bar{X}_{T2})$: is the estimate for the test treatment drug averaged over both periods,

$\bar{X}_R = \frac{1}{2}(\bar{X}_{R1} + \bar{X}_{R2})$: is the estimate for the reference treatment drug averaged over both periods.

Bioequivalence studies usually start with the same number of study subjects, n_T and n_R, in sequence 1 and sequence 2, but the number that completes the study may not be equal. According to Chow and Liu (2000) it should be noted that $\bar{X}_T$ and $\bar{X}_R$ are the equivalent to $\bar{Y}_T$ and $\bar{Y}_R$, least square means for the test and reference formulations respectively.

According to Rani and Pargal (2004) an ANOVA gives the mean square error (MSE), which is needed to obtain an unbiased estimate of ${\sigma_d}^2$ (the pooled sample variance of period differences from both sequences). MSE, in ANOVA Table 2.1 is used when calculating the confidence intervals.

Table 2.1 ANOVA Table

Sources of variation	**Degree of freedom (DF)**	**Sum of squares (SS)**	**Mean sum of squares (MS)**	**F Statistic**
Treatment	t^a*-1*	*SST*	*MST*	*MST/MSE*
Subject	n^b*-1*	*SSS*	*MSS*	*MSS/MSE*
Period	*t-1*	*SSP*	*MSP*	*MSP/MSE*
Error	*(t-1)(n-2)*	*SSE*	*MSE*	
Total	*tn-1*			

The 90% confidence interval approach discussed on untransformed data also applies when data is transformed. For example, using the pharmacokinetic parameter, *AUC*, if transformed it becomes *lnAUC*.

The mean after the logarithmic transformation becomes:

$$\bar{I}n\ AUC = \sum_{i=1}^{2} \frac{In\ AUC_{it}}{n} \qquad (2.7.2.1)$$

where the subscript *t* represents the test treatment formulation of the *AUC* for the i^{th} study subject and *n* is the sample size.

The geometric mean is obtained by transforming the mean back to original data.

Geometric mean$= e^{(ln\ AUC_t)}$ (2.7.2.2)

The confidence interval approach, first suggested by Westlake (1976), states that bioequivalence may be concluded if a $(1-2\alpha)\times100\%$ confidence interval for the difference $\mu_T - \mu_R$ or ratio $\frac{\mu_T}{\mu_R}$ is within an acceptance range (α is usually set at 0.05). If the ±20 rule (for original data) is used, this means that the confidence interval for the difference of the means of the two treatment formulations must be between -0.2 and 0.2. Likewise, the confidence interval for the ratio of means must be totally included in the interval 0.8 to 1.2 or 80% to 120%. The lower bound (LB) for the confidence interval is 0.8 and 1.2 is the upper bound (UB) for the confidence interval. Average bioequivalence is achieved if the confidence interval is totally included in the equivalence interval 0.80 to 1.25 or 80% to 125% for the logarithmic transformed data, where 0.8 is the LB and 1.25 is the UB. There are several methods which can be used for calculating confidence intervals. Researchers on bioequivalence studies suggest that among the several available methods, the most appropriate one for a particular research project should be selected.

2.7.2.1 Classic Confidence Interval of the Difference of the Means

Let $\bar{Y}_T$ and $\bar{Y}_R$ be the respective least squares means for the test and reference formulations. The classic or shortest $(1-2\alpha)\times$ 100% confidence interval can be obtained based on the following *t* statistic:

$$T=\frac{(\bar{Y}_T-\bar{Y}_R)-(\mu_T-\mu_R)}{\hat{\sigma}_d\sqrt{\frac{1}{n_T}+\frac{1}{n_R}}}$$

where n_T and n_R are the number of subjects in sequences 1 and 2, respectively and $\hat{\sigma}_d$ the mean square error is obtained from $\hat{\sigma}^2{}_d$, the pooled sample variance of period differences from sequences. L_1 and U_1 are the lower and upper confidence interval bounds for the interval of the difference of the means.

$$L_1 = (\bar{Y}_T - \bar{Y}_R) - (t_{\alpha;n_T + n_R - 2})\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}}$$

and (2.7.2.1.1)

$$U_1 = (\bar{Y}_T - \bar{Y}_R) + (t_{\alpha;n_T + n_R - 2})\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}}.$$

2.7.2.2 Classic Confidence Interval of the Ratio of the Means

A confidence interval for the ratio: $\frac{\mu_T}{\mu_R}$ may be computed from the confidence interval on the difference of the means of two treatments by dividing by $\bar{Y}_R$ to obtain the formula:

$$L_2 = (L_1 / \bar{Y}_R + 1) \times 100\%$$

and (2.7.2.2.1)

$$U_2 = (U_1 / \bar{Y}_R + 1) \times 100\%.$$

L_2 and U_2 are the lower and upper confidence interval bounds for the ratio of the means for the test and reference treatment formulation.

2.7.2.3 Westlake's Symmetric Confidence Interval of the Difference of the Means

Compute values of k_1 and k_2 so that, according to Chow and Liu (2000) the classic confidence interval derived from an unpaired two sample *t* statistic in equation 2.7.2.1.1 is: $k_1 < T < k_2$ where *k* is the upper α^{th} percentile of a central *t* distribution with $(n_T + n_R - 2)$ degrees of freedom. In general, a $(1-2\alpha)\times 100\%$ confidence interval for the difference $\mu_T - \mu_R$ can be expressed as:
$k_1 < T < k_2$, where k_1 and k_2 are chosen so that the probability from k_2 to k_1 based on a central *t* distribution with $(n_1 + n_2 - 2)$ degrees of freedom is $(1-2\alpha)$, that is,

$$1 - 2\alpha = \int_{k_2}^{k_1} T dt$$

$$\Delta = k_1\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}} - (\bar{Y}_T - \bar{Y}_R)$$

and (2.7.2.3.1)

$$\Delta = -k_2\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}} + (\bar{Y}_T - \bar{Y}_R).$$

Finally, conclude bioequivalence if

$$|\Delta| < 0.2\mu_R \quad (2.7.2.3.2)$$

2.7.2.4 Westlake's Symmetric Confidence Interval of the Ratio of the Means

A confidence interval for the ratio may be computed from the confidence interval on the difference of means using the formula

$$L_4 = (-|\Delta|/\ \bar{Y}_R\ + 1) \times 100\%$$

and (2.7.2.4.1)

$$U_4 = (|\Delta|/\ \bar{Y}_R\ + 1) \times 100\%.$$

2.7.2.5 Confidence Interval of the Ratio of Means Based on Fieller's Theorem

The (1−2α)×100% confidence limits for $\delta = \frac{\mu_T}{\mu_R}$ are the roots of the quadratic equation

$$(\bar{Y}_T - \delta\bar{Y}_R)^2\ - (t_{\alpha;\, n_1 + n_2 - 2})^2\omega\ (S_{TT} - 2\delta S_{TR} + \delta^2 S_{RR}\ ) = 0 \quad (2.7.2.5.1)$$

where;

$$\omega = \frac{1}{4}\left(\frac{1}{n_T} + \frac{1}{n_R}\right) \quad (2.7.2.5.2)$$

$$S_{RR} = \frac{1}{\left(\frac{1}{n_1} - \frac{1}{n_2} - 2\right)}\left[\sum_{i=1}^{n_1}(Y_{i11} - \bar{Y}_{.11})^2 + \sum_{i=1}^{n_2}(Y_{i22} - \bar{Y}_{.22})^2\right] \quad (2.7.2.5.3)$$

$$S_{TT} = \frac{1}{\left(\frac{1}{n_1} - \frac{1}{n_2} - 2\right)}\left[\sum_{i=1}^{n_1}(Y_{i21} - \bar{Y}_{.21})^2 + \sum_{i=1}^{n_2}(Y_{i12} - \bar{Y}_{.12})^2\right] \quad (2.7.2.5.4)$$

$$S_{TR} = \frac{1}{\left(\frac{1}{n_1} - \frac{1}{n_2} - 2\right)}\left[\sum_{i=1}^{n_1}(Y_{i11} - \bar{Y}_{.11})(Y_{i21} - \bar{Y}_{.21}) + \sum_{i=1}^{n_2}(Y_{i12} - \bar{Y}_{.12})(Y_{i22} - \bar{Y}_{.22})\right] \quad (2.7.2.5.5)$$

Additionally, in order for the roots of the quadratic equation to be real, rational and positive, the above values must satisfy the conditions

$$\frac{\bar{Y}_R}{\sqrt{\omega S_{RR}}} > t_{\alpha; n_T + n_R - 2}$$

and (2.7.2.5.6)

$$\frac{\bar{Y}_T}{\sqrt{\omega S_{TT}}} > t_{\alpha;\, n_T + n_R - 2}.$$

In this research report the classic or shortest confidence interval (C.I.) approach will be used to assess ABE. The choice of using this C.I. approach instead of others, discussed above, is based on the fact that this approach is globally applied and accepted on evaluating ABE. Classic C.I. has a weakness of that the required level of equivalence to conclude ABE, 0.8 to 1.25, may not be achieved if the within subject variability is large. To overcome this weakness, bootstrapping, as described in Sections 3.12 and 4.7 can be used. Limitations of other C.I. approaches seem to be more than the weakness of the classic C.I. method.

Westlake' symmetric confidence interval approach is a modification of the classic C.I. Many researchers criticize Westlake's C.I. because they are symmetric about μ_R instead of $\bar{Y}_T - \bar{Y}_R$ and the tail probabilities for the hypotheses are not symmetric, they divert from being two sided, becoming a one sided hypothesis as $\mu_T - \mu_R$ or the standard error increases. Westlake can be used for decision making not necessarily for estimation and hypothesis testing. Confidence interval based on Fieller's theorem is very attractive since variability of the treatment formulations is accounted for. However, Fieller's method's limitations are that it is derived using very mild assumptions of normality and is an approximation, not an exact procedure since the method is developed by substituting an estimate of the intra subject correlation (Chow and Liu, 2000).

2.7.3 Interval Hypothesis Testing Approach

In order to investigate average bioequivalence the student *t*-test or Fisher test are not recommended since the hypothesis tested is not the same. For the classical student *t*-test as well as the Fisher test (ANOVA), the hypothesis tested is:

$H_0: \mu_T = \mu_R$ bioequivalence

$H_1: \mu_T \neq \mu_R$ bioinequivalence (2.7.3.1)

Where, μ_T represent the population mean for the test treatment and μ_R stands for the reference treatment formulation. While, the hypothesis for a bioequivalence study is:

Additive bioequivalence test of hypotheses

$H_0: \mu_T - \mu_R < \Delta_1$ or $\mu_T - \mu_R > \Delta_2$ bioinequivalence

$H_1: \Delta_1 \leq \mu_T - \mu_R \leq \Delta_2$ bioequivalence (2.7.3.2)

$(\Delta_1 ; \Delta_2)$ represents absolute equivalence interval.

Multiplicative bioequivalence test of hypotheses

$H_0: \frac{\mu_T}{\mu_R} < \Delta_1$ or $\frac{\mu_T}{\mu_R} > \Delta_2$ bioinequivalence

$H_1 : \Delta_1 \leq \frac{\mu_T}{\mu_R} \leq \Delta_2$ bioequivalence (2.7.3.3)

$(\Delta_1 ; \Delta_2)$ is the relative equivalence interval where, $0<\Delta_1<1<\Delta_2$ for example, 0.8 to 1.25.

Multiplicative hypotheses for the bioequivalence test become additive after a log transformation

$H_0: ln\mu_T - ln\mu_R < ln\Delta_1$ or $ln\mu_T - ln\mu_R > ln\Delta_2$

$H_1: ln\Delta_1 \leq ln\mu_T - ln\mu_R \leq ln\Delta_2$ (2.7.3.4)

2.7.3.1 Schuirmann's Two One-Sided Tests Procedure

Schuirmann (1987) introduced the idea of using an interval hypothesis to evaluate ABE, applying the following null and alternative hypotheses:

$$H_0: \mu_T - \mu_R \leq \theta_L \quad or \quad \mu_T - \mu_R \geq \theta_L$$
$$\theta_L < \mu_T - \mu_R < \theta_U \qquad (2.7.3.1.1)$$

where θ_L and θ_U are limits selected to satisfy bioequivalence. The hypotheses 2.7.3.1.1 can be separated to form two one-sided tests below:

$$H_{01}: \mu_T - \mu_R \leq \theta_L \text{ versus } H_{\alpha 1}: \mu_T - \mu_R > \theta_L$$
$$H_{02}: \mu_T - \mu_R \geq \theta_U \text{ versus } H_{\alpha 2}: \mu_T - \mu_R < \theta_L \qquad (2.7.3.1.2)$$

The first hypothesis test whether the treatment effect is too low and the second tests whether the treatment effect is too high. If both null hypotheses are rejected, it can be concluded that the test treatment is average bioequivalent compared to the reference treatment formulation by using the Schuirmann's two one-sided tests procedures. Jones and Kenward (2003) states that the method gets its name, TOST, because the process of deciding if the 90% confidence interval lies within the acceptance limits is the same as rejecting both of the following one-sided hypotheses in equation 2.7.3.1.3 at 5% level of significance.

$$H_{01}: \mu_T - \mu_R \leq -ln1.25 \text{ versus } H_{02}: \mu_T - \mu_R \geq ln1.25 \qquad (2.7.3.1.3)$$

If both tests are rejected, the conclusion of bioequivalence is made at α significance level. That is, conclude that μ_T and μ_R are average bioequivalent at the α significance level if

$$T_L = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_L}{\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}}} > t_{\alpha;\, n_T + n_R - 2}$$

and (2.7.3.1.4)

$$T_U = \frac{(\bar{Y}_T - \bar{Y}_R) - \theta_U}{\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}}} > -t_{\alpha;\, n_T + n_R - 2}$$

If the hypothesis being tested leads to the same conclusion, parametric tests will be preferred, but if the conclusion is not the same, nonparametric tests are then used.

2.7.3.2 Anderson and Hauck's Test

Unlike Schuirmann's tests procedures, Anderson and Hauck (1984) suggested a one sided hypothesis interval approach for concluding ABE between two treatment formulations. The significance level of the Anderson and Hauck test is given by:

$$\alpha = \Pr(|t_{AH}| - \hat{\delta}) - \Pr(-|t_{AH}| - \hat{\delta}) \quad (2.7.3.2.1)$$

where;

$$\Pr(x) = \int_{-\infty}^{x} T_{n_T + n_R - 2}\, dt \quad (2.7.3.2.2)$$

$$\hat{\delta} = \frac{\theta_U - \theta_L}{\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}}} \quad (2.7.3.2.3)$$

$$t_{AH} = \frac{(\bar{Y}_T - \bar{Y}_R) - (\theta_U + \theta_L)/2}{\hat{\sigma}_d\sqrt{\frac{1}{n_T} + \frac{1}{n_R}}} \quad (2.7.3.2.4)$$

Anderson and Hauck's test (1984) has a higher probability of concluding ABE as compared to the C.I. approaches discussed in Sections 2.7.2.2, 2.7.2.3, 2.7.2.4 and 2.7.2.5 and is also powerful contrasted to Schuirmann's TOST procedures. However, Anderson and Hauck's weaknesses are mainly on that the actual significance level (α) may be bigger than the nominal significance level and this test may also conclude ABE even when variability is very large due to the fact that its rejection region is open ended. As a result of the serious limitations of

Anderson and Hauck's test, Schuirmann's TOST procedures are widely used (Chow and Liu, 2000).

Bayesian methods can also be used to assess average bioequivalence. The main Bayesian methods according to Chow and Liu (2000), Jones and Kenward (2003) include the Rodda and Davis method as well as the Mandallaz and Mau's method.

2.7.4 Wilcoxon Mann-Whitney Two One-Sided Test Statistics

Chow and Liu (2000) indicated that because the 2×2 crossover design consists of two sequences, that is, RT and TR, a distribution free sum of the ranks test can be applied directly to the two one-sided tests procedure. The Wilcoxon Mann-Whitney two one-sided test statistics is applied when assessing average bioequivalence for the pharmacokinetic parameter T_{max} because this parameter does not exhibit normality hence a distribution free test is applied. The Wilcoxon Mann-Whitney two one-sided test statistics is as follows:

Let: $\theta = \mu_T - \mu_R$ (2.7.4.1)

The following hypotheses are tested:

$H_{01}: \theta_L^* \leq 0 \quad \text{versus} \quad H_{a1}: \theta_L^* > 0$

and (2.7.4.2)

$H_{02}: \theta_U^* \geq 0 \quad \text{versus} \quad H_{a2}: \theta_U^* < 0.$

where

$\theta_L^* = \theta - \theta_L$ (2.7.4.3)

$\theta_U^* = \theta - \theta_U$ (2.7.4.4)

Let $b_{hik} = \begin{cases} \mathrm{d}_{ik} - \theta_h & h = \mathrm{L}, \mathrm{U}, \text{for subjects in sequence 1} \\ \mathrm{d}_{ik} & \text{for subjects in sequence 2} \end{cases}$ (2.7.4.5)

Where $d_{ik} = \frac{1}{2}\,(\mathrm{Y}_{i2k} - \mathrm{Y}_{i1k}), \quad i = 1,2,\ldots,n_k; k = 1,2.$ (2.7.4.6)

When carryover effects are absent, the expected value and variance of $\mathrm{b_{hik}}$ is given by:

$$\mathrm{E}(b_{hik}) = \begin{cases} \frac{1}{2}\,[(\mathrm{P}_2 - \mathrm{P}_1) + (\theta - 2\theta_h)] & \text{for } k = 1 \\ \frac{1}{2}\,[(\mathrm{P}_2 - \mathrm{P}_1) - \theta] & \text{for } k = 2 \end{cases} \quad (2.7.4.7)$$

where $h = \mathrm{L}, \mathrm{U}, i = 1,2, \ldots, \mathrm{n_k}$ and $k = 1,2$ and $\mathrm{V}(b_{hik}) = \mathrm{V}(d_{ik})$

$$= \sigma_{\mathrm{d}}^2$$

$$= \frac{\sigma_{\mathrm{e}}^2}{2} \qquad (2.7.4.8)$$

Therefore, $E(b_{hi1}) - E(b_{hi2}) = (\theta - \theta_h)$

$$= \theta_h^* \qquad (2.7.4.9)$$

Define R_L to be the value obtained by adding ranks of the responses for study subjects in sequence 1, given as:

$$R_L = \sum_{i=1}^{n_1} R\ (b_{Li1}) \qquad (2.7.4.10)$$

Therefore, the Wilcoxon Mann-Whitney test statistic for H_{01} is:

$$W_L = R_L - \frac{n_1(n_1+1)}{2} \qquad (2.7.4.11)$$

H_{01} is rejected if $W_L > w(1-\alpha)$, (2.7.4.12)

where w$(1-\alpha)$ is the $(1-\alpha)th$ quantile of the distribution of W_L.

Similarly, for the second set of hypotheses:

H_{02}: $\theta_U^* \geq 0$ versus H_{02}: $\theta_U^* < 0$ we reject H_{02} if

$$W_U = R_U - \frac{n_1(n_1+1)}{2} < w\ (\alpha) \qquad (2.7.4.13)$$

Where R_U is the sum of the ranks of $\{b_{hi2}\}$ for subjects in sequence 1. Hence, average bioequivalence can be concluded if H_{01} and H_{02} are both rejected, that is:

$W_L > w\ (1-\alpha)$ and $W_U < w(\alpha)$ (2.7.4.14)

The expected values and variances for W_L and W_U under the null hypotheses H_{01} and H_{02}, when there are no ties, are given by:

$$E(W_L) = E(W_U) = \frac{n_1 n_2}{2} \qquad (2.7.4.15)$$

$$V(W_L) = V(W_U) = \frac{1}{12}\ n_1 n_2\ (n_1 + n_2 + 1) \qquad (2.7.4.16)$$

If some ranks are equal, assign the mean of the ranks to compute W_L and W_U. In this case, however, the expected values and variances of W_L and W_U become:

$$E(W_L) = E(W_U) = \frac{n_1 n_2}{2} \qquad (2.7.4.17)$$

$$V(W_L) = V(W_U) = \frac{1}{12}\ n_1 n_2\ (n_1 + n_2 + 1 - Q) \quad (2.7.4.18)$$

where

$$Q = \frac{1}{(n_1 + n_2)(n_1 + n_2 - 1)} \sum_{V=1}^{q} (r_V^3 - r_V) \quad (2.7.4.19)$$

Where q is the number of groups with the same rank values and r_V is the magnitude of ranks with the same value in group *V*. Suppose all observations are different, $q = n_1 + n_2\ ; r_V = 1$ for V=1,2,…, n and Q=0, equation 2.7.4.19 reduces to equations 2.7.4.17 and 2.7.4.18.

Since W_L and W_U are symmetric about their mean, $\frac{n_1 n_2}{2}$, we have $w\ (1 - \alpha) = n_1 n_2 - w(\alpha)$.

When $n_1 + n_2$, is a large sample size (say, $n_1 + n_2 > 40$) and the ratio n_1 and n_2 is close to $\frac{1}{2}$, the distribution becomes approximately normal (symmetric):

$$W_L > w\ (1 - \alpha) \text{ and } W_U < w(\alpha) \quad (2.7.4.20)$$

for average bioequivalence testing, that is, we may conclude bioequivalence if

$$Z_L > z(\alpha) \text{ and } Z_U < -z(\alpha) \quad (2.7.4.21)$$

where z(α) is the α^{th} quantile of a symmetric distribution, and

$$Z_L = \frac{W_L - E(W_L)}{\sqrt{V(W_L)}} = \frac{R_L - [\frac{n_1(n_1+n_2+1)}{2}]}{\sqrt{\frac{1}{12}\ n_1 n_2 (n_1 + n_2 + 1)}} \quad (2.7.4.22)$$

$$Z_U = \frac{W_U - E(W_U)}{\sqrt{V(W_U)}} = \frac{R_U - [\frac{n_1(n_1+n_2+1)}{2}]}{\sqrt{\frac{1}{12}\ n_1 n_2 (n_1 + n_2 + 1)}} \quad (2.7.4.23)$$

Note that the expected values and variances in Z_L and Z_U should be replaced with that given in equations 2.7.4.17 and 2.7.4.18 if ties exist.

2.7.5 Population Bioequivalence

According to Chen and Lesko (2001), population bioequivalence puts emphasis mainly on prescribability (pre-approval), that is, can a patient be safely and effectively started on some other drug formulation.

The population bioequivalence (PBE) criterion is:

Reference-scaled criterion

$$\theta_{PBE} = \frac{(\mu_T-\mu_R)2+(\sigma_T^2-\sigma_R^2)}{\sigma_R^2} \leq \theta_p, \text{ for } \sigma_R^2 > \sigma_O^2 \qquad (2.7.5.1)$$

Constant-scaled criterion

$$\theta_{PBE} = \frac{(\mu_T-\mu_R)2+(\sigma_T^2-\sigma_R^2)}{\sigma_O^2} \leq \theta_p, \text{ for } \sigma_R^2 \leq \sigma_O^2 \qquad (2.7.5.2)$$

where;

θ_p = Population bioequivalence limit,
μ_T = Mean of the test treatment formulation,
μ_R = Mean of the reference treatment formulation,
σ_T^2 = The total variance of the test treatment formulation,
σ_R^2 = The total variance of the reference treatment formulation,
σ_O^2 = Specified threshold value of the total variance.

The reference-scaled criterion is to be applied when the total variance, σ_R^2 is greater than an FDA (2001) specified threshold value, σ_O^2 =0.04. Otherwise, the constant-scaled criterion is used. Population bioequivalence is concluded if it can be demonstrated that $\theta_{PBE} \leq \theta_p$ where the recommended value for θ_p =1.7448, found as follows:

$$\theta_p = \frac{Average\ bioequivalence\ limit + variance\ allowance}{Scaled\ variance}$$

$$= \frac{\ln(1{,}25)^2 + (\sigma_T^2-\sigma_R^2)}{\sigma_O^2}$$

$$= \frac{\ln(1{,}25)^2 + 0.02}{0{,}04}$$

$$= 1.7448$$

The FDA (2001) guidance document uses the notation $\sigma_{TT}^2 = \sigma_T^2$ and $\sigma_{TR}^2 = \sigma_R^2$ for the total variances of the respective formulations. The hypothesis tested is:

H_0^{PBE}:$\theta_{PBE} \gg \theta_P$ versus H_1^{PBE}:$\theta_{PBE} < \theta_P$. (2.7.5.3)

PBE is concluded if H_0^{PBE} is rejected.

2.7.6 Individual Bioequivalence

Chen (2007) states that the FDA started investigations as from 1992 to find out whether average bioequivalence is adequate for the assessment of comparisons between test and reference treatment formulations. In 1999 the question of whether average bioequivalence is sufficient or not, was posed to the generic drugs Advisory Committee Meeting to deliberate on the issue, check whether average bioequivalence has limitations and motivate the need for assessing individual bioequivalence as prerequisite for approval of generic drugs. Subsequently groups were formed through the Centre for Drug Evaluation and Research of the FDA from the academics, industry and regulatory authority to explore the issue of individual bioequivalence. The underlying questions to address this issue were asked:

- Is there a need to change the basis of bioequivalence criteria from average bioequivalence to individual bioequivalence?
- What are the desirable properties of bioequivalence criteria if we are not satisfied with the current criteria based on average bioequivalence?
- What are the general approaches in developing the criteria of individual bioequivalence if this concept is chosen for bioequivalence assessment? (FDA, 1999).

Switchability is the main thrust of individual bioequivalence. It looks at the possibility of safely and effectively switching a patient from the original marketed formulation to another. Individual bioequivalence evaluates three components:

- comparison of means,

- compares subject by formulation interaction,
- compares within subject variances.

The key issues involved in the formulation of the criteria for individual bioequivalence are the individual therapeutic window and intra subject variability. An individual therapeutic window can be viewed as the distance between the minimum effective exposure and the maximum tolerable exposure of a drug, whereas the intra subject variability can be viewed as the distribution of the individual's responses to the drug. In theory, if a drug product given is both efficacious and safe to a patient, the distribution of his or her responses should fall within his or her own therapeutic window. When a test product is introduced that is interchangeable with the reference product, the distribution of this patient's responses from the test product must also fall within his or her therapeutic window (Chen, 2007).

Therapeutic window is defined as an interval of bioavailability metric or pharmacokinetic response such as $AUC_{0-\infty}$ in which the drug is efficacious and safe (Chow and Liu, 2000). The lower and upper limits of a therapeutic window are known as: the minimally effective level (MEL) and the maximally tolerated level (MTL) respectively. If the difference between the MTL and MEL is large, the therapeutic window is wide while if the difference is small, the therapeutic window is narrow.

The individual bioequivalence criterion (IBE) is:

$$\sigma_D^2 = \sigma_{BT}^2 + \sigma_{BR}^2 - 2\rho\,\sigma_{BT}\sigma_{BR}$$
$$= (\sigma_{BT} - \sigma_{BR})^2 + 2(1-\rho)\,\sigma_{BT}\sigma_{BR} \qquad (2.7.6.1)$$

Reference-scaled criterion

$$\theta_{IBE} = \frac{(\mu_T-\mu_R)2+\sigma_D^2+(\sigma_{WT}^2-\sigma_{WR}^2)}{\sigma_{WR}^2} \leq \theta_1, \text{ for } \sigma_{WR}^2 > \sigma_{WO}^2 \qquad (2.7.6.2)$$

Constant-scaled criterion

$$\theta_{IBE} = \frac{(\mu_T-\mu_R)2+\sigma_D^2+(\sigma^2_{WT}-\sigma^2_{WR})}{\sigma^2_{WO}} \leq \theta_1, \text{ for } \sigma^2_{WR} \leq \sigma^2_{WO} \qquad (2.7.6.3)$$

Here σ^2_{WO} is a constant (pre-set value), of which the FDA (2001) recommended value for σ^2_{WO} =0.04

IBE is concluded if $\theta_{IBE} \leq \theta_1$. Currently, the FDA (2001) recommended value for θ_1 =2.2448 when σ^2_D =0.02 and θ_1 =2.4948 when σ^2_D =0.03.

$$\theta_1 = \frac{Average\ bioequivalence + variance\ allowance}{Scaled\ variance}$$

$$= \frac{\ln(1.25)^2 + \sigma^2_D + (\sigma^2_{WT}-\sigma^2_{WR})}{\sigma^2_{WO}}$$

$$= \frac{\ln(1.25)^2 + 0.02 + 0.02}{0.04}$$

$$= 2.2448.$$

σ^2_O or $\sigma^2_{WO}$the pre-specified threshold value for both PBE and IBE is 0.04.

where;

θ_1 =Individual bioequivalence limit,

μ_T =Mean of the test treatment formulation,

μ_R =Mean of the reference treatment formulation,

σ^2_D =Subject by formulation interaction component of the variance,

($\sigma^2_{BT} - \sigma^2_{BR}$) are between subject variances of the test and reference formulations respectively,

σ^2_{WT} =Within subject variance of the test treatment formulation,

σ^2_{WR} =Within subject variance of the reference treatment formulation.

The denominator of σ^2_{WR} provides a scaling factor FDA (2001) suggested that the reference scaled individual bioequivalence be applied when the estimated

magnitude of σ_{WR}^2 exceeds a pre-set value of the variance (σ_{WO}^2) :$(\sigma_{WR}^2 > \sigma_{WO}^2)$, otherwise, a constant-scaled criterion is used.

The hypothesis tested is:

H_0^{IBE}:$\theta_{IBE} \gg \theta_I$ versus H_1^{IBE}:$\theta_{IBE} < \theta_I$ (2.7.6.4)

IBE is concluded if H_0^{IBE} is rejected.

2.8 Comparison of Average Bioequivalence, Population Bioequivalence and Individual Bioequivalence

Endrenyi, Amidon, Midha and Skelly (1998) observe that it is not easy to use the models for individual bioequivalence, they are more complicated and have more parameters than those applied for average bioequivalence. More questions arose as well investigating whether individual bioequivalence is really necessary or not:

- Has average bioequivalence failed, that is, have there been documented problems observed following the substitution of bioequivalence?
- Is there evidence that subject by formulation interactions are important?
- What populations, for example, healthy volunteers or patients, are appropriate for the evaluation of IBE and, in particular, for the study of the subject by formulation interaction?
- Are comparisons of within (intra) subject and intra formulation variations relevant? (Endrenyi *et al*, 1998).

All these questions were dealt with by groups formed at the Centre for Drug Evaluation and Research. One such group comprises of Chen, Patnaik, Hauck, Schuirmann, Hyslop, and William (2000) who worked on individual bioequivalence as well as developing the statistical methods to assess bioavailability measures. Individual bioequivalence has an advantage over average and population bioequivalence because it allows assessment of subjects by formulation interaction and compares population averages as well as variances between the test and reference treatment drugs. Population bioequivalence

accounts for average bioequivalence only, variances are not considered, hence as a result, there is need for individual bioequivalence to complement population bioequivalence. The need for individual bioequivalence is further enhanced by that pharmacokinetic responses for test and reference products differ among individuals. Meredith (2003) is among such researchers who are argued that population bioequivalence should go hand in hand with individual bioequivalence. He states that bioequivalence is achieved by certifying two treatment drugs as ABE, but, there are concerns on the use of ABE only since it does not effectively evaluate treatment formulations with a narrow or wide therapeutic interval or high intra subject or inter subject variability. To alleviate the fears observed using ABE, the use of IBE is suggested.

Drug interchangeability, which can be classified as either drug prescribability or drug switchability differentiates between population and individual bioequivalence. The treatment formulation a physician chooses to prescribe to a patient compared to the other available drugs is known as drug prescribability. Drug switchability is a case where a patient is switched from one treatment formulation to another, ensuring that the concentration of the drug has been toned to a steady, efficacious and safe level as compared to the first one. Switchability, an attribute obtained through individual bioequivalence, is synonymous to a guarantee that once drug products have been certified as bioequivalent, they can then be used interchangeably in the target population, hence the seriousness by all stakeholders currently in pursuing the blending of population and individual bioequivalence. Prescribability is assured by population bioequivalence through comparing the population means and variance components between the test and the reference treatments. The concept of individual bioequivalence is really important because it ensures that, for a certain period, as deemed sufficient by regulatory authorities (Chen *et al,* 2000). The bioavailability or bioequivalence measure for an individual remains unchanged in order to ascertain safety and efficacy of that individual (FDA, 2001).

Average bioequivalence is compared and contrasted with individual bioequivalence, of which the observations are:

- Average bioequivalence is used by the FDA to approve the marketing of thousands of generic drugs;
- For a few other drugs, in vivo studies are waived and market access is granted based on in vitro studies;
- Large amount of empirical evidence suggests that generic drugs are used regularly without problems of safety or efficacy.

However, on the other hand, the realisations were that individual bioequivalence:

- Has been proposed by the FDA (2001) as an improvement on the study design, in formativeness and method of analysis of bioequivalence studies;
- Controversial topic with many debates and public discussion;
- Has not been universally accepted (Chen and Lesko, 2001).

The other drawback of IBE is that it allows patients to be switched from a reference treatment formulation to its corresponding test treatment. However, it falls short regarding switching a patient from one test treatment to another generic drug formulation, a trend happening in practice (Patterson, 2010).

To show that test and reference formulations are average bioequivalent it is only necessary to show that the ratio of the means AUC and C_{max} for test and reference treatment formulations are within the accepted bioequivalence interval as per a regulatory authority specifications. It is possible for one drug to be much more variable than the other, yet similar in terms of the means, therefore, ABE has some limitation since it does not consider the variability. Population bioequivalence is a measure which combines the mean and variance of the drug products. Individual bioequivalence has more information to enable the approval of generic drugs as bioequivalent though there are also challenges on using IBE models. The way forward which seem to be a general consensus among researchers on bioequivalence, is that the concepts involved in population bioequivalence need to be integrated with individual bioequivalence so as to be

thorough on ascertaining generic drugs as being bioequivalent (McCarthy and Guo, 2008).

The next chapter looks at the criteria used for selecting study subjects for a bioequivalence study, assuming human beings are used. The crossover design is discussed and implemented using data obtained from the study subjects (sheep). Data from sheep dosed using the RT and TR was assessed for ABE and PBE using various statistical techniques.

Chapter 3

Methodology

3.1 Introduction

This chapter focusses on the methods of how the test treatment and reference treatment formulations will be compared in terms of their bioavailability. Methods include in vitro studies, comparative clinical studies, pharmacodynamics studies and pharmacokinetic studies. Pharmacokinetic studies are preferred as compared to the other methods of assessing bioequivalence and justification is given as to why they are chosen in a bioequivalence study.

Bioequivalence regulatory authorities prescribe standard requirements for candidates to be used in the study, tables with such requirements are available in this chapter. A 2×2 crossover design is used once study subjects have been identified and enrolled to find pharmacokinetic parameters. Some study subjects, with very low or extremely high bioavailability, known as outliers are discussed in this chapter, including the various methods for detecting them. Methods of statistical analysis are classified into parametric and nonparametric, where the former method is applied on the pharmacokinetic parameters: AUC_{0-t}, $AUC_{0-\infty}$ and C_{max} but T_{max} utilizes the latter method. This chapter also discusses the power of a test, bootstrapping and effect sizes.

According to the FDA (2003) and Mastan *et al.* (2011), bioequivalence between a test drug and reference drug can be assessed through mainly four ways, namely:

- in vitro studies,
- comparative clinical studies,
- pharmacodynamic studies,
- pharmacokinetic studies.

The methods of assessing bioequivalence, mentioned above, are applied depending on the type of treatment formulation or method of administering the drug(s). They can be used in isolation or as combinations. Bioequivalence is well

established through pharmacokinetic studies. Pharmacodynamics studies are not recommended where pharmacokinetic methods can be used. Clinical studies can be utilized for situations where other methods for assessing bioequivalence cannot be used (Wang and Bakhai, 2006). In vitro studies are more appropriate for immediate release (IR) solid oral dosage forms (Polli, 2008). The FDA (2003) guidance for bioequivalence has a clause, 21 *CFR* 320.24, which clearly indicates that clinical studies need to be avoided if possible. These guidelines also indicate that bioequivalence can be tested by in vitro or in vivo studies given immediate release solid oral dosage forms approved after 1962 and for bio-problem treatment formulations approved before 1962. Pharmacokinetic studies are usually used or preferred in assessing bioequivalence because the parameters obtained enable the researchers to measure bioavailability. The measurements taken should indicate the concentration of the drug in the body, that is, the amount of the treatment drug in a given volume of blood. If for some reasons, measurements of the drug cannot be extracted through the blood, then measurements can be achieved through urinary excretion.

3.2 Study Subjects for a Bioequivalence Study

Study subjects which are preferred for a bioequivalence study should possess attributes like:

- The study subjects for population bioequivalence are chosen in such a manner that there is maximum variability among them but differences between pharmaceutical products should be noticeable without much difficulty;
- Healthy volunteers of either sex, in the range 18 to 55 years old are normally preferred to be in the study;
- Women who still have a potential of giving birth are screened on an individual basis;
- The weight of study subjects is also considered before enrolling into the bioequivalence study, the recommended weight should be proportional to the body mass index. Non-smokers, who have never consumed alcohol, are preferred;

- Candidates for study are thoroughly screened and undergo a medical examination before they can be enrolled (FDA, 2003).

Below are Tables 3.1, 3.2 and 3.3, showing regulatory requirements (human beings) for demographics, sample size and whether study subjects need to be fed or fast for different regulatory bodies from Tamboli, Todkar and Sayyad (2010) contrasted with FDA (2003) and MCC (2003) bioequivalence requirements. Regulatory requirements in tables below were included in this research report to discuss what is needed if human beings were to be used in a bioequivalence study.

Table 3.1 Regulatory Requirements for Demographics

Regulatory Authority	Age	Gender	BMI (kg/m²)
India	Greater than or equal to 18 years, but if the treatment formulation is for use by the elderly, the majority of the study subjects should be 60 years of age or older.	Gender selection should be consistent and depend on usage and the safety criteria.	Not specified.
USA	18 years or older.	Male or female.	Not specified.
Europe	18 years or older.	Male or female.	18,5 to 30kg/m²
Canada	18 to 55 years.	Males or female.	Ratio of height/weight should be within 15% of the normal interval of healthy volunteers.
ASEAN	18 to 55 years.	Male or female.	18 to 25kg/m²
South Africa	18 to 55 years.	Male or female.	Recommended BMI or within 15% of the normal body mass.
Brazil (ANVISA)	18 to 50 years.	Male or female.	Within 15% of the accepted interval.

Table 3.2 Regulatory Requirements for Sample Size

Regulatory Authority	Minimum	Sample size specifications
India	16 years or more.	The number of subjects required for a study should be statistically significant and should be large enough to allow for possible drop outs from the study.
USA	12 years	The total number of subjects in the study should provide sufficient power to conclude bioequivalence.
Europe	12 years or more.	The sample size calculation method determines the sample size.
Brazil	12 years or more.	The number of subjects required is determined by: • The error variance linked with the main characteristic to be studied as estimated from a pilot experiment obtained from previous studies or from published data. • The desired level of significance. • The expected deviation from the reference treatment formulation compatible with bioequivalence. • The required power.
South Africa	12 years or more subjects for immediate release and 20 subjects for modified release oral dosage forms	The number of subjects should be justified on the basis of providing at least 80% power of achieving the acceptance criteria or appropriate equations should be used to calculate the sample size.

Table 3.3 Fasting and FED Study Requirements

Regulatory Authority	Fasting requirements
India	Should fast for 10 hours or more at night followed by 4 hours of fasting after taking a treatment formulation. If dosing is done many times at night, fasting should be 2 hours before and 2 hours after dosing.
Europe and Brazil	Greater than or equal to 8 hours before and greater than or equal to 4 hours after dosing.
USA and Canada	Should be greater than or equal to 10 hours and continued for at least 4 hours after dosing.
ASEAN	Greater than or equal to 8 hours before taking the treatment formulation.
South Africa	The same time period of fasting should be maintained before and after dosing.

All of these criteria above were relevant if the data that were used for bioequivalence analyses were for human trials, but they were from an animal trial. Animals used in bioequivalence studies must be healthy and from a homogeneous group (age, breed, sex, weight, hormonal and nutritional status, level of production among other aspects). When it is difficult to conserve homogeneity of all animals within a study (for example, sheep), it would be acceptable to use non-homogeneous stock provided that animals in each treatment group were carefully matched for characteristics including age, weight and sex. Selected animals must be representative of the target population for which the product is intended. Group size: for ethical and economic reasons, the appropriate number of animals should be carefully estimated; it depends on several factors including variance of the response, differences in the two treatment formulations and level of rejection of the hypothesis (The European Agency for the Evaluation of Medicinal Products, 2001).

3.3 Crossover Design

The two sequences, two periods, crossover study design is recommended by many researchers on assessing bioequivalence using pharmacokinetics parameters. In

this research report, generic and brand name products are administered to study subjects (experimental units) with the objective of trying to establish differences (if any) due to availability of generic and brand name products on different study subjects. A crossover study seeks to investigate the response of individual study subjects to two different treatments (Jones and Kenward, 2003).

Sequence	**Period 1**	**Washout**	**Period 2**
1	R		T
2	T		R

Figure 3.1 2×2 Crossover Design

A crossover design is a design used when collecting blood samples from study subjects such that pharmacokinetic parameters can be obtained. In a crossover model, study subjects are randomized into two groups then given a sequence of treatments at uniform intervals of time (periods). In the first sequence, each study subject receives, say, a test treatment in period 1, then a reference treatment in period 2 and this arrangement is reversed in the second sequence, reference treatment is given followed by the test treatment.

Wang and Bakhai (2006) illustrated the 2×2 crossover design diagrammatically as shown below:

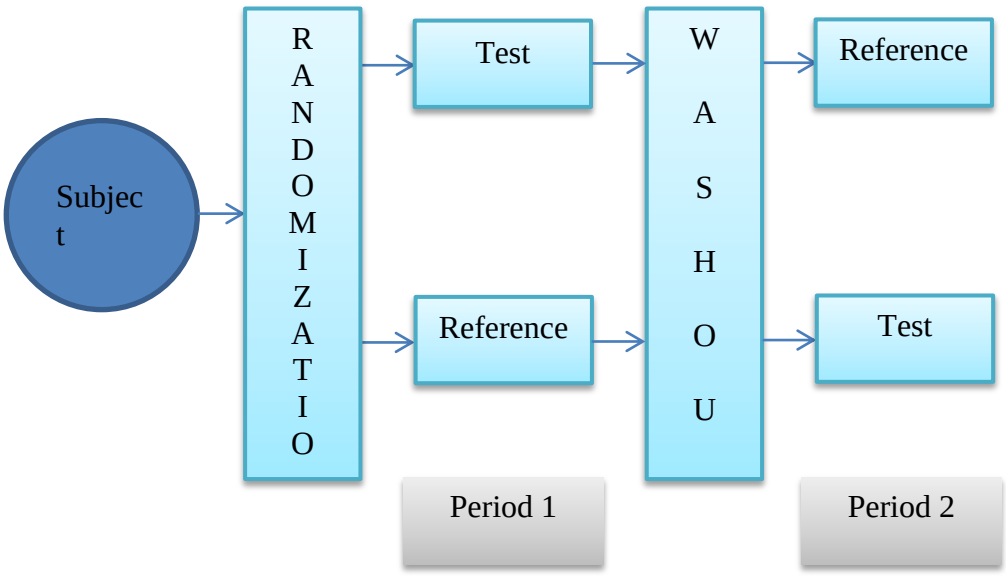


Figure 3.2 2×2 Crossover Design

A model for the 2×2 crossover design from Chow and Liu (2000) is given as:

$$Y_{ijk}=\mu+S_{ik}+P_j+F_{(j,k)}+C_{(j-1,k)}+\varepsilon_{ijk} \quad (3.3.1)$$

where;

Y_{ijk} represents the response, for example, *AUC* of the i^{th} subject in the k^{th} sequence at the j^{th} period,

μ =The overall mean of the model,

S_{ik} =The random effect of the i^{th} subject in the k^{th} sequence where i=1,2,…,g,

P_j =The fixed effect of the j^{th} period, where j=1,…,p and $\sum_j P_j$=0,

$F_{(j,k)}$ =The direct fixed effect of the formulation in the k^{th} sequence which is administered at the j^{th} period and $\sum F_{(j,k)}$ =0,

$C_{(j-1,k)}$ =The fixed first order carryover effect of the formulation in the k^{th} sequence which is administered at the $(j-1)^{th}$ period where $C_{(0,k)}$ =0 and $\sum C_{(j-1,k)}$ =0,

ε_{ijk} =The random error in observing Y_{ijk} .

It is assumed that S_{ik} are independently and identically distributed (iid) with mean 0 and variance σ_s^2, and ε_{ijk} are independently distributed with mean 0 and variance σ_t^2, where t = 1,2,…,L (the number of formulations to be compared). S_{ik} and ε_{ijk} are assumed to be mutually independent. The estimate of σ_s^2 is usually used to explain the inter-subject variability, and the estimates of σ_t^2 are used to assess the intra subject variabilities for the n^{th} formulation.

3.3.1 Advantages of a Crossover Design

The crossover designs are preferred as compared to other experimental designs because they are not influenced much by physiological variables. Greenblatt *et al.* (1998) states that physiological variables such as age, gender, body height, ethnicity, hepatic and renal disease might affect drug disposition. Using crossover designs, therefore, enables the researcher(s) to find the difference between the

drug concentrations of a generic and a brand name drug due to formulations of the drug.

The crossover design is viewed favourably because of its advantages:

- Each subject receives both treatment formulations hence a comparison can be made within subjects;
- It eliminates the inter subject variability from comparison between formulations;
- With a proper randomization of subjects to the sequences of treatment formulations, the crossover design provides the best unbiased estimates for the differences (or ratios) of the means (Jones and Kenward, 2003; FDA, 1992 and Chow and Liu, 2000).

Each trial subject acts as his own control since both drug formulations are administered per study subject. The crossover design is thus very useful because both treatments, generic and brand name, are given to the same subject. Drug concentration measurements are taken from each study subjects for both treatment formulations such that bioequivalence can be assessed (Galpin, 2007).

3.3.2 Disadvantage of a Crossover Design

The crossover design is a powerful study design where comparison is made within study subjects though it has a disadvantage of that it is not suitable for treatment formulations with long half-life and the effect of a treatment given in one period, called a carryover effect, might still be present at the start of the following period.

3.4 Carryover Effect

There are cases where it is not possible to totally eliminate the effect of a previous treatment resulting in a carryover effect. A carryover effect is a situation where some amount of a treatment (drug) from one period will still be available when blood samples are taken on the next period. Senn (2001) indicates that a carryover effect arises if a treatment (drug) administered in one period continue to affect the study subject in subsequent period(s). He goes on to state that a carryover effect

introduces bias to estimates of the amounts of a particular drug in blood samples (generic or brand name), because the assumption will be that the effect of one treatment is being observed yet two or more drugs will be still within the study subject(s).

The carryover effect arises due to a short interval of time allowed to pass before administering the next drug formulation to a study subject. Other causes of a carryover effect could be as a result of:

- Drugs which have a long half-life, though with an inadequate washout period;
- Trial subjects with poor metabolism rates, hence taking a long time to eliminate the drug in the body;
- Random occurrence (Patterson, 2010).

In bioequivalence studies a long washout period is allowed between dosing periods such that the carryover effect is eliminated. The washout period is the interval between dosing periods. In a crossover design, the washout period should be long enough such that the effect of one treatment formulation administered at one period is not available at the next period of dosing (Chow and Liu, 2000). Jones and Kenward (2003) indicated that a washout period, between dosing periods, should be at least five half-lives so as to ensure that the carryover effects are eliminated completely. As a consequence of having a long washout period, there is no need to test for a differential carryover effect (FDA, 2001).

The 2×2 crossover model without the carryover effect becomes:

$$Y_{ijk} = \mu + S_{ik} + P_j + F_{(j,k)} + \varepsilon_{ijk} \qquad (3.4.1)$$

where the components of this model are defined in equation 3.3.1.

The crossover design, Figure 3.2, was used at the University of Pretoria at the Directorate of Veterinary Pharmacology where sheep were randomized then dosed with the reference treatment formulation (an antibiotic) followed by the test treatment formulation after an adequate washout period was allowed before the

next dosing period to obtain the bioequivalence data. The model 3.3.1 was used in this study when assessing ABE and PBE.

3.5 Pharmacokinetic Parameters

The pharmacokinetic parameters obtained after collecting the blood samples from study subjects are C_{max}: the rate (but also extent) of absorption. Estimated directly from the data, where C_{max} =max(C_0,C_1,C_2,…C_k).

T_{max}: is the corresponding time point estimate at which C_{max} occurs.

Chow and Liu (2000) and Wang and Bakhai (2006) explained how to calculate the *AUC*: the extent of drug absorption, obtained as follows:

$$AUC_{0-t} = \sum_{i=1}^{t}(\frac{C_i + C_{i-1}}{2})(T_i - T_{i-1}) \quad (3.5.1)$$

$$AUC_{0-\infty} = AUC_{0-t} + AUC_{t-\infty}$$

$$= AUC_{0-t} + \frac{C_{last}}{\lambda} \quad (3.5.2)$$

$$\lambda = \text{Slope} * -2.303 \quad (3.5.3)$$

Slope is obtained from the plots of ln(Concentration) in Appendix 4, at the time points where the elimination of the drugs become approximately linear.

Calculations for finding the AUC in this study were done from Appendix 1 to Appendix 5.

3.6 Parametric versus Nonparametric Tests

Parametric and nonparametric tests will be used in this research study to analyse the figures obtained after collecting blood samples from the study subjects in order to determine the pharmacokinetics parameters. Given a data set, a researcher needs to make a choice as to whether use parametric or nonparametric tests so that reasonable decisions based on the data obtained can be taken. The selection of the tests to use is guided by the nature of the data, as indicated earlier. If the data follows a normal distribution or approximately so, the parametric tests are recommended in order to get best results but if the data does not follow a normal distribution, nonparametric tests are appropriate. There are quite a number of ways which can be used to determine whether the data set is normal or not, common approaches of checking normality include: plotting a histogram, finding the coefficient of skewness or kurtosis. Kurtosis is a measure which indicates

whether the data is peaked or flat relative to a normal distribution. It is easy to visualize whether the data follows a normal distribution or is skewed if a histogram is plotted. When determining normality by skewness, the coefficient of skewness is used. The coefficient of skewness for a normal distribution is equal to zero. Data sets with a coefficient of skewness greater or less than zero need to be transformed to try and reduce skewness: The transformations which can be applied on data include the log transformation, square root, square, inverse, bucketing and many more. However, in this report, the choice of tests was as per the FDA (2001) recommendations since most of the research done on bioequivalence is per criteria established by the FDA.

3.6.1 Parametric Tests

Concordet (2004) points out that a statistical property of the distribution of data is that all data can be completely described by a finite number of parameters. Guassian (Normal) distribution, N($\mu; \sigma^2$), is an example of a parametric distribution. Parametric distributions assumptions are:

- homoscedasticity,
- independence,
- normality.

Homoscedasticity refers to a situation where the variance of the dependent variable is constant, that is, it does not vary with independent variables which are: the drug formulation, study subject and period.
Independence is a situation whereby the random variables which feature in the bioequivalence analysis are independent.

Normality in a distribution is achieved if the pharmacokinetic parameters used in the bioequivalence analysis are normally distributed.

For bioequivalence studies the pharmacokinetic parameters which usually utilise the parametric tests are the AUC and C_{max}.

ABE and PBE were used to test for bioequivalence between the two drugs for the sheep, on the pharmacokinetic parameters.

3.6.2 Nonparametric Tests

Nonparametric tests are applied for situations where the distribution of data is not defined by a finite number of parameters. It can only be defined by its shape, number of modes, regularity and other features. It should be noted that for nonparametric distributions, the number of parameters used to estimate the distribution with *n* data increases with *n*. A known statistical distribution cannot be linked to nonparametric tests, that is, we cannot say that this distribution of data follows a Normal, Exponential, Binomial, Poisson, Geometric or some other of the common statistical distributions. The nonparametric tests are usually less powerful, since it is more difficult to show bioequivalence even for cases where there is bioequivalence, as compared to their parametric counterparts but nonetheless there are situations which require nonparametric methods (where the distribution is not normal, as pointed out earlier). For bioequivalence studies, T_{max}, is analysed by using nonparametric methods (FDA, 2001).

Nonparametric tests are a branch of statistical inference (decision making) whereby the estimation or hypothesis tested is not necessarily about a population or sample statistic and distribution assumptions are not used as compared to the parametric tests. According to Galpin (2007) nonparametric tests are appropriate for a data set which is severely skewed whereby even transforming the data to approximate normality is a futile exercise, hence normality of the data cannot be assumed and there is no indication as to which distribution the data can follow. When a data set is normally distributed, the mean is equal to the median and we use the mean as our measure of centre. However, if the data set is skewed, then the median is a much better measure of centre. Therefore, just like the *Z* (normal distribution), *t* (Student test) and *F* (Analysis of Variance) tests made inferences about the mean(s) on parametric tests, nonparametric tests make inferences about the population median(s).

Parametric tests were used in this study when assessing ABE and PBE for the pharmacokinetic parameters *AUC* and C_{max} because the histogram for the data, Appendix 7, Figure 1 and Appendix 8, Figure 1, exhibits a distribution which is almost normal. Non-parametric test, Wilcoxon Mann-Whitney test statistics were used in this study to evaluate ABE and PBE for the parameter T_{max} because the data is not normally distributed.

3.7 Outliers

In a bioequivalence study there are observations obtained from the drug concentration time profiles which are extremely high or low as compared to the rest, such observations are called outliers. Outliers are probable as a result of an error in typing data or a data value(s) obtained mistakenly from a study subject(s) which is/are not supposed to be included in the research for a variety of reasons. Outliers are defined by Patterson (2010) in terms of the residuals in the model, whereby a data value that is too large or small cannot be fitted into the model, hence becomes an outlier. There are four different types of outliers:

- Observations which arise unexpectedly in concentration time profiles;
- Observations which are very huge or minute in one of the treatment formulations;
- A large significant difference, extremely large or small, in bioavailability when the generic drug is contrasted to the reference formulation;
- Observations which exhibit an unusually high or low concentration of drug in a study subject in both treatment formulations (Chow and Liu, 2000).

Causes of outliers are varied, though basically common according to many authors on bioequivalence. Frequent changes in laboratory tests may result in some observations becoming too high or low, hence becoming outliers. Human error on measuring the concentration of blood can also cause an outlier (Karasoy and Daghan, 2012).

Other causes of outliers may be attributed to:

- Defects in a drug, treatment formulation may be broken (though coated) or having a wrong dosage;
- Sudden changes in a study subject, probable someone starts vomiting or develops diarrhoea;
- Laboratory error, say, a wrong drug is prescribed;
- A study subject can have an unusual reaction to a treatment formulation or both drugs resulting in a subject by formulation interaction which was not expected (Schall, Ring and Endrenyi, 2010).

There are tests which can be used to detect outlying subjects. Chow and Liu (2000), Ramsoy and Elkum (2004) and Karasoy and Daghan (2012) discussed methods of identifying outliers:

3.7.1 Likelihood Distance Test (LD)

The LD was developed on the basis that the period and treatment formulation effects do not exist on the 2×2 crossover model 3.4.1. The model 3.4.1 with the period and treatment formulation equal to zero reduces to:

$$y_{ij} = \mu + S_i + e_{ij,} \quad j = 1, \ldots., f, \; i = 1, \ldots., n. \qquad (3.7.1.1)$$

The parameters of interest in model 3.7.1.1 are: μ, σ_s^2 and σ_e^2. Let $\theta = (\theta_1, \theta_2, \theta_3)^t$, where $\theta_1=\mu$, $\theta_2 = \sigma_e^2$, and $\theta_3 = \sigma_e^2 + f\sigma_s^2$. The log-likelihood function for the model 3.7.1.1 is:

$$L(\theta)= \frac{-fn}{2}\log 2\pi - \frac{n}{2}\log(\theta_3\theta_2^{f-1}) - \frac{1}{2\theta_2}\sum_{i=1}^{n}\sum_{j=1}^{f}\left(Y_{ij} - \theta_1\right)^2 - \frac{f}{2}\left(\frac{1}{\theta_3} - \frac{1}{\theta_2}\right)\sum_{i=1}^{n}\left(\bar{Y}_i - \theta_1\right)^2 \qquad (3.7.1.2)$$

The maximum likelihood estimators (MLEs) of the parameters θ_1, θ_2 and θ_3 derived by maximizing $L(\theta)$ given that $\theta_2=\theta_3$ are:

$$\hat{\theta}_1=\bar{Y} = \frac{1}{nf}\sum_{i=1}^{n}\sum_{j=1}^{f} Y_{ij} \qquad (3.7.1.3)$$

$$\hat{\theta}_2=\frac{1}{n(f-1)}\sum_{j=1}^{f}(Y_{ij}-\theta_1)^2 \quad (3.7.1.4)$$

$$\hat{\theta}_3=\frac{\mathrm{f}}{\mathrm{n}}\sum_{\mathrm{i}}(\overline{Y}_{\mathrm{i}}-\overline{Y})^2 \quad (3.7.1.5)$$

The LD statistic becomes:

$$LD_i(\hat{\theta}) = 2\left[L(\hat{\theta}) - L(\hat{\theta}_{(i)})\right] \quad (3.7.1.6)$$

where $\hat{\theta}_{(i)}$ is the MLE of θ after deleting the i^{th} subject. It can be shown that as *n* approaches infinity $LD_i(\hat{\theta})$ is asymptotically distributed as a chi-square with three degrees of freedom. The i^{th} subject is considered to be an outlier if $\mathrm{LD_i}(\hat{\theta}) > \chi^2_{3,}(\alpha)$ where $\chi^2_{3,}(\alpha)$ is the α^{th} upper percentile of a central chi-square distribution with three degrees of freedom.

3.7.2 Estimates Distance Test (ED)

ED is a method of detecting outliers and was developed based on the difference of the parameter estimates obtained after deleting the i^{th} subject, slightly different from the LD which is based on the difference in the log-likelihood function after deleting the i^{th} subject.

$$ED_i(\hat{\theta}) = f^2\,(\hat{\theta}-\hat{\theta}_{-i})^t\hat{\sum}^{-1}(\hat{\theta}-\hat{\theta}_{-i}) \quad (3.7.2.1)$$

Equation 3.7.2.1 is the ED statistic where $\hat{\sum}^{-1}$ is the MLE of the variance matrix below

$$\sum=\begin{bmatrix}\theta_3/n & 0 & 0\\ 0 & 2\theta_2^2/(n-1) & 0\\ 0 & 0 & 2\theta_3^2\end{bmatrix} \quad (3.7.2.2))$$

The $ED_i(\hat{\theta})$ statistic was shown to be asymptotically distributed as a chi-square with three degrees of freedom by Chow and Tse (1990). The i^{th} subject is identified as an outlier by the ED test if $ED_i(\hat{\theta}) > \chi^2_{3,}(\alpha)$.

3.7.3 Hotelling T^2 Test (HT)

This is an outlier detection procedure based upon the order statistics of the two-sample Hotelling T^2 statistic. This procedure was proposed by Liu and Weng (1991) according to Ramsoy and Elkum (2004). HT is based on the assumption that there are no period effects and the compound symmetry covariance structure for f responses observed on the i^{th} subject is relaxed such that it becomes spherical. These assumptions enable HT to identify only one outlier. In practice, pharmacokinetic parameters can have more than one outlier. To detect outlying subjects, let $Y_i = (Y_{i1,\dots,}Y_{if})^t$ be the vector of the response variables observed on the i^{th} subject. Define

$$D_i^2 = (Y_i - \overline{Y})^t A^{-1} (Y_i - \overline{Y}) \qquad (3.7.3.1)$$

where $\overline{Y}$ and A are the sample mean and the matrix of the sums of squares and cross-products computed from $Y_1, \dots, Y_n$. The HT statistic for the i^{th} subject is:

$$T_i^2 = \frac{(n-2)D_i^2}{(\frac{n-1}{n} - D_i^2)} \qquad (3.7.3.2)$$

To detect whether subject i is an outlier or not, we compare the value obtained from equation 3.7.3.2 to the critical value. The i^{th} subject is identified to be an outlier if at α =0.05 level of significance:

$$T_i^2 = \max_j \left(T_j^2\right) > T_{(\alpha)}^2 \qquad (3.7.3.3)$$

3.7.4 Mean-Shift Test (MS)

The mean-shift test was developed by Wang and Chow (2003) based on the likelihood function. The mean shift for the t^{th} subject's response on the j^{th} treatment formulation is:

$$Y_{ij} = \mu + S_i + F_j + \Lambda_j \delta_{it} + e_{ij} \qquad (3.7.4.1)$$

Two quantities developed from the model 3.7.4.1 are:

$$T_{1n} = \frac{(e_i - \bar{e}_t \boldsymbol{I})'(e_i - \bar{e}_t \boldsymbol{I})}{\sum_s (e_s - \bar{e}_s \boldsymbol{I})'(e_s - \bar{e}_s \boldsymbol{I})} \quad (3.7.4.2)$$

and

$$T_{2n} = \frac{\bar{e}_t{}^2}{\sum_s \bar{e}_t{}^2}. \quad (3.7.4.3)$$

where in equations 3.7.4.2 and 3.7.4.3, e_t is the vector of residuals for the t^{th} subject, $\bar{e}_t$ stands for the mean of e_t and $\boldsymbol{I}$ is a vector whose members are 1. Combining equations 3.7.4.2 and 3.7.4.3 the MS statistic is:

$$D_t = nT_{1n} + nT_{2n} \quad (3.7.4.4)$$

The statistic D_t is used to test whether the t^{th} subject is an outlier or not and is distributed as the sum of two independent beta random variables:

$$D_t : nBeta[1/2, (n-1)/2] + Beta[1/2, (n-1)/2] \quad (3.7.4.5)$$

The t^{th} subject is detected as an outlier if D_t is greater than the T^2 order statistic at α =0.05 level of significance for a given sample size.

3.7.5 Residuals Test

The residuals test utilizes the means of treatment formulations to detect outlying subjects. Given $Y_{.j}$ the j^{th} treatment formulation mean, the resulting studentized residuals, r_{ij} where i=1,2,…,n and j = 1,2,…,f. The test compares the maximum values of the residuals with the critical values. The i^{th} subject is identified to be an outlier if the maximum residual: $r_{ij} > T_\alpha\,(n)$.

The Cook's distance, likelihood distance test (LD), Figure 4.1 and Figure 4.2 in Chapter 4 is used to detect the outlying observations for the pharmacokinetic parameters $AUC_{0-\infty}$ and C_{max} respectively. The likelihood distance test could be obtained using the model 3.7.1.1, and equations 3.7.1.2 and 3.7.1.6 though in this study the model in equation 3.4.1 is applied to get LD, where the pharmacokinetic parameters $AUC_{0-\infty}$ and C_{max} for both untransformed and transformed data are

subjected to ANOVA in which the variance is partitioned into components due to animals (sheep), periods, sequence and treatment effects.

3.8 Outliers simulation

Karasoy and Daghan (2012) conducted a simulation bioequivalence study using twenty three study subjects to detect outlying observations. The same observations were identified to be outliers by the LD, ED, HT MS tests. Confidence intervals obtained for the pharmacokinetic parameter *AUC* including the outlier and without the outlier were both totally included in the C.I. acceptance range, 0.8 to 1.25. The outlier changed the ratio of means but equivalence limits were not affected much, Table 3.4.

Table 3.4 Outliers Ratio of Means and Confidence Intervals

AUC	90% C.I. with outliers		90% C.I. without outliers	
Ratio of means	Lower limit	Upper limit	Lower limit	Upper limit
	0.9581	0.9989	0.9815	0.9975
Ratio of means	0.979		0.990	

Ramsoy and Elkum (2004) also conducted a simulation study to identify outliers using LD, ED, HT and MS. The same outlying subjects were detected by all the methods, similar to the study done by Karasoy and Daghan (2012). Ramsoy and Elkum (2004) further contrasted the tests checking if any was better. They observed that all tests performed well, though ED outperformed other methods where there is only one outlier.

3.9 Power of a Test

Statistical power analysis is performed at the planning stage of a bioequivalence study since it assists in improving the chances of achieving efficient results when assessing average bioequivalence. Appropriate sample size for a study is determined by power analysis as well as the research objective, design, data

analysis method, type 1 error, variability and effect size. In a bioequivalence study, sample size, refers to the number of subjects or volunteers participating in the study. Chow and Liu (2000) point out that appropriate sample size for a study is chosen depending on a power function of the statistic for the hypothesis of equality between treatment formulation effects.

The statistical power of a test is defined, in general, as the probability of rejecting the null (research) hypothesis when it is actually false. (reject H_0 when H_0 is false). Ahmad and Yahya (2015) states that in terms of efficacy for a drug, power of a test refers to the probability of correctly concluding that a drug is effective when in fact it is. With reference to a 2×2 crossover design used in bioequivalence studies, power implies not mistakenly concluding that the two treatment formulations are bioequivalent when there are not. Important aspects to be known about the power of a test:

- As the sample size increase, the power of the test also increases. That is, by taking larger samples, we improve our ability to find a difference in means of the treatment formulations, if they really exist;
- As the population variances decrease, the power of the test increases. It should, however, be noted that the researcher has no control over the variances;
- As the difference in the means: $\mu_1 - \mu_2$ increases, the power also increases. The difference in the means of the two treatment formulations can also not be controlled by the researcher (Elsayir, 2012).

Factors that affect the power of a test are mainly, the sample size, significance test and effect size:

- If the sample size is large, the power of a test increases;
- If the level of significance is high, the power of a test also becomes high because the acceptance region is reduced;

- If there is a large difference between the true value of the mean of the treatment formulations and the hypothesized value of the mean, the power of a test becomes high (Chow and Liu, 2000).

The power of a test is determined from the outcomes of a hypothesis test. When a hypothesis test is performed, there are four possible outcomes. Two of the four outcomes are correct while the other two are wrong. The conclusions which are not correct are known as the type 1 and type 2 errors, also sometimes referred to as consumer's risk and producer's risk respectively. Four conclusions on hypothesis testing are:

H_0 is true and your test leads you fail to reject H_0: correct decision;
H_0 is true but your test leads you to reject H_0: wrong decision (Type 1 error);
H_0 is false but your test leads you to fail to reject H_0: wrong decision (Type 2 error);
H_0 is false and your test leads you to reject H_0: correct decision (Crawshaw and Chambers, 2001).

The four conclusions for a hypothesis test can be depicted on a table, Table 3.5:

Table 3.5 Classical Type 1 and Type 2 errors

		Test decision	
		Fail to reject H_0	Reject H_0
Actual situation	H_0 is true	Correct decision	Type 1 error
	H_0 is false	Type 2 error	Correct decision

The power of a test is defined as the probability of correctly rejecting H_0 when H_0 is false. The choice of the null hypothesis is based on the seriousness of the errors which can be committed. Chow and Liu (2000) illustrate the relationship between type 1 and type 2 errors for a bioequivalence trial as in Table 3.6:

Table 3.6 Bioequivalence Type 1 and Type 2 errors

Decision	True State H_0	
	Bioinequivalent	Bioequivalent
Bioinequivalent	Right decision	Type 2 error
Bioequivalent (reject H_0)	Type 1 error	Right decision

A type 1 error, also referred to as the consumer's risk in bioequivalence terms, is more serious contrasted to a type 2 error, the producer's risk. Hence to control both, that is, having the errors remaining as minimal as possible, the significance level is controlled at an acceptable level, while there should be a way to try to minimize the type 2 error, usually done by choosing an appropriate sample size (Chow and Liu, 2000).

3.10 Simulation of Power

The TOST procedure is normally preferred for power simulation in bioequivalence studies. It is difficult to calculate the exact power for TOST since the relevant formulae require some complex numerical integration. Simulation is used as an alternative to computations given that an algebraic approach cannot be found. The intention is to find a sample size which will achieve a certain desired power. Power should generally be at least 80%. Simulating power empirically yields an approximate power estimate which is nearly accurate because a large number of data sets are generated as per specifications of the distribution used for the power analysis.

A bioequivalence study was conducted with the objective of finding a sample size necessary to achieve at least 95% power. Two different treatment formulations were used with a geometric mean ratio of 1.1 and variance 0.1003. Hypothesis tested is:

H_0:$\tau_B - \tau_A \ll ln(0.8)$ or $\tau_B - \tau_A \gg ln(1.25)$ versus

H_1:$ln(0.8) < \tau_B - \tau_A < ln(1.25)$ (3.10.1)

Table 3.7 Power and sample size SAS 9.4 output

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Lower Equivalence Bound	-0.22314
Upper Equivalence Bound	0.223144
Alpha	0.05
Mean Difference	0.09531
Standard Deviation	0.223942
Nominal Power	0.95
Computed N per Group	
Actual Power	N per Group
0.952	68

The SAS 9.4 output, Table 3.7, indicates that 68 study subjects will be required to obtain an actual power of 95.2% (Sun, 2010).

3.11 Power curve

Table 3.8 Power curve and sample size SAS 9.4 output

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Null Difference	0
Alpha	0.05
Mean Difference	4
Standard Deviation	3
Number of Sides	2

Computed N per Group			
Index	Nominal Power	Actual Power	N per Group
1	0.80	0.805	10
2	0.90	0.903	13
3	0.95	0.954	16

The three groups in Table 3.8 results in different power values per given sample size. As the sample size increase in a particular group, the power also increases. Sample group size 10, power 80.5%, sample group size 13, power increases to 90.3% and a further power increase to 95.2% is achieved when the sample group size is 16. The power curve, Figure 3.3 illustrates visually the increase in power as the sample group size increases (Plets and Strominger, 2013).

This study had two groups of 7 sheep in sequence 1 and the other 7 were in sequence 2. Comparing the sample size of the groups in this study and the groups in the simulation study, Figure 3.3, it was observed that more sheep were needed in this study to obtain a power of at least 80%.

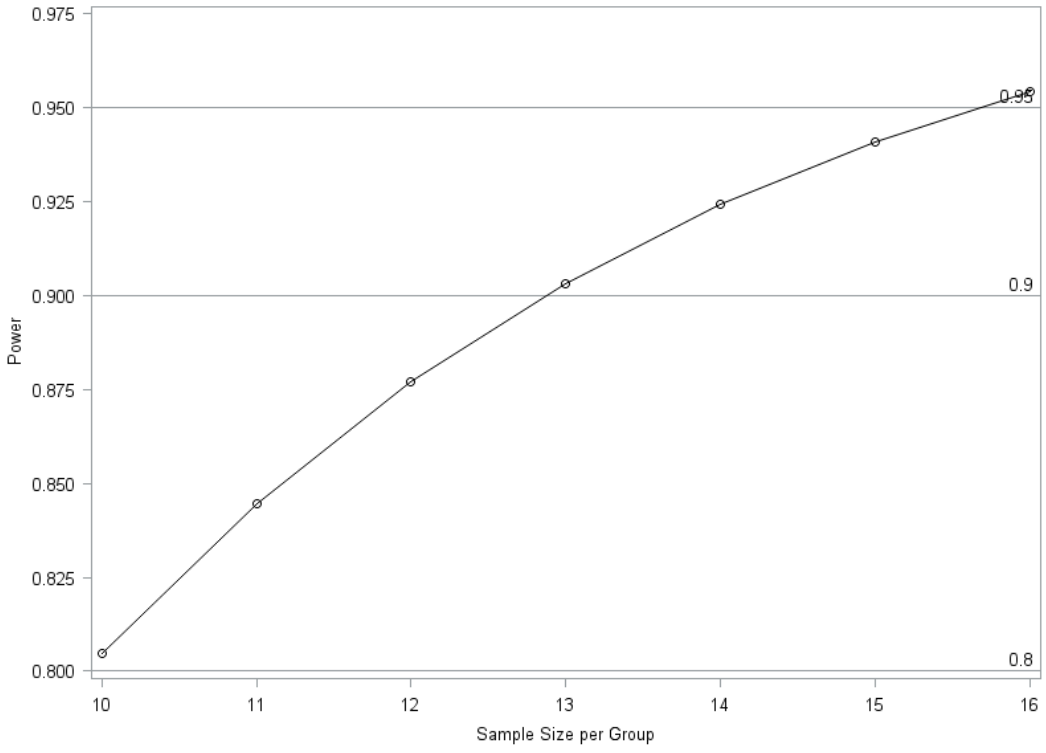


Figure 3.3 Power curve and sample sizes

3.12 Bootstrapping

Bootstrapping is a technique which enables researchers to do statistical inference without checking any model assumptions and without any sampling distribution. Bootstrapping is a nonparametric method which refers to sampling with replacement from the original data. It depends solely on the original data, thus avoiding as many assumptions compared to parametric methods (Moony and Duval, 2012; Schmidheiny, 2012 and Rochowicz, 2011). Bootstrapping is also defined as a resampling method which is effectively performed using a computer for estimating sample statistics such as the measures of spread, confidence intervals, hypothesis testing and other statistical properties. The number of bootstrap samples depends on what you like to do. If bootstrapping to estimate the standard error, twenty five to two hundred resamples should be sufficient but for all other applications, bootstrap samples should be more than one thousand. For a situation where there is a small number of data values in a given sample, bootstrapping technique can be used to increase the sample size. Population parameter estimates can then be obtained from this large sample which could have been formed through bootstrapping. As a result, aspects of the population such as skewness, kurtosis and percentiles can be checked (Efron and Tibshirani, 1993).

Boostrapping is a powerful statistical technique and a very useful tool used when the distribution of a statistic is unknown, very complex or when the sample size is small and asymptotic distribution assumptions such as normality may not be appropriate. The bootstrap method in short is:

- Resample a given data set with replacement a specified number of times, where each, bootstrap sample has the same number of data points as the number in the original sample though some of the values of the original sample might not appear, some might appear once, twice or thrice;
- Calculate a statistic of interest for each of the bootstrap samples;
- The distribution of the statistic from the bootstrap samples can then be used to obtain estimated standard errors, create confidence intervals, and to perform hypothesis testing with the statistic (Rice and Thornotn, 2013).

Bootstrapping statistics actually enables a researcher to analyse any distribution and make inferences. Bootstrapping liberates researchers from being limited to doing statistical analysis using known sampling distributions only by using computers. Bootstrapping as a computer intensive method, can be easily performed nowadays because modern computers have fast processors and are within the reach of researchers (Chernick, 2008). Besides bootstrapping as a resampling method, there are other techniques available in literature, which include the jacknife, cross–validation, random subsampling and permutation procedures. Permutation tests unlike, bootstrap, are procedures whereby resampling is done though not randomly, instead, it considers all possible permutations (arrangements) of the sample (Moony and Duval, 2012).

Bootstrapping has pros and cons:

Advantages

- Checking assumptions of a distribution is not required.
- Can be used for cases where permutation tests fail because bootstrap requires very minimum assumptions.
- A large sample size is obtained.

Disadvantages

- Efficient computers in terms of speed are needed.
- Randomness when sampling with replacement should be understood.
- Bootstrap is not exact.
- Large sample sizes must be generated though for large samples, permutation tests perform better than bootstrap (Moony and Duval, 2012).

Bootstrapping was done in this study, Appendix 9, to illustrate the fact that as the sample size is increased statistical inference can be made without checking any model assumptions.

3.13 Assessing Average, Population and Individual Bioequivalence

3.13.1 Average Bioequivalence

The 90% confidence interval for the ratio of means of the parameter C_{max} and *AUC* should lie within the interval 0.80 to 1.25, the bioequivalence acceptance range. The acceptance interval may need to be changed especially in cases where the drugs have a narrow therapeutic range. In such cases a wide acceptance range of 0.90 to 1.111 should be applied (Schall and Endrenyi, 2010).

Quite a number of regulatory authorities use the above interval to certify a generic formulation as average bioequivalent compared to a reference formulation. Below is a table, Table 3.9 showing different regulatory authorities accepted intervals for the pharmacokinetic parameters *AUC* and C_{max}.

It should be noted that in South Africa the average bioequivalence acceptance interval for the pharmacokinetic parameter C_{max} of 0.75 to 1.33 differs with most regulatory authorities. However for formulations which have a narrow therapeutic range, the average bioequivalence acceptance range is the same as the regulatory authorities on Table 3.9, for most regulatory authorities (countries), that is, the interval from 0.8 to 1.25 (Galgate, Jamdade, Aute, Chaudhari, 2013).

Table 3.9 Regulatory Acceptance Criteria for Bioequivalence (Tamboli *et al*, pp.91).

Regulatory Authority	90% confidence interval on log transformed data		
	C_{max}	AUC_{0-t}	$AUC_{0-\infty}$
India	80-125	80-125	80-125
USA	80-125	80-125	80-125
Europe and Australia	80-125	80-125	Not applicable
South Africa	75-133 and 80-125 (for narrow therapeutic range)	80-125	Not applicable
ASEAN	80-125	80-125	80-125

Decision rules based on confidence intervals and the acceptance range on certifying drug formulations as average bioequivalent or not:

- C.I. entirely outside the acceptance range. Average bioinequivalence proven;
- C.I. overlaps the acceptance range. Average bioequivalence not proven;
- C.I. lies entirely within the acceptance range. Average bioequivalence proven (Schutz, 2013).

3.13.2 Population Bioequivalence

To establish population bioequivalence, unbiased estimators obtained using the method-of-moments (MM) are used on logarithmic transformed values of C_{max} and *AUC*. Population bioequivalence is achieved if the 90% upper confidence bound calculated after finding unbiased estimators is less than or equal to the population bioequivalence limit, θ_p =1.7448, as recommended by the FDA (2001).

3.13.3 Individual Bioequivalence

Individual Bioequivalence focuses on estimation of the mean difference between test and reference formulation, the subject by formulation interaction variance and the within subject variance for each of the formulations. To estimate the components of IBE the FDA (2001) recommends the method-of-moments approach by Chinchilli and Esinhart (1996). It is also indicated that the restricted maximum likelihood (REML) method is useful to estimate mean difference and variances when subjects with missing data are included in the statistical analysis. A 90% upper confidence bound for individual bioequivalence is calculated and then compared to the individual bioequivalence limit. The FDA (2001) recommended value for θ_1 =2.2448 when σ_D^2 =0.02 and θ_1 =2.4948 when σ_D^2 =0.03. σ_D^2 represents the subject by formulation variance component when calculating θ_1. Individual bioequivalence is achieved if the 90% upper confidence bound is less than or equal to the bioequivalence limit.

Individual bioequivalence is necessary in bioequivalence studies since it takes into account between subject variation, within subject variability and subject by formulation interaction, hence having the property of switchability, whereby patients can be changed from one drug formulation to the other, a property not found when assessing bioequivalence using ABE and PBE. However, in this research report, IBE cannot be calculated since the data used is a from a standard 2×2 crossover design which is not appropriate for assessing IBE (FDA, 1999; 2001; Chow and Liu, 2000 and Jones and Kenward, 2003). The FDA (1999) guidance for industry and FDA (2001) guidance on statistical approaches recommends the higher order designs such as: [RTR, TRT] or [RTRT, TRTR] for the evaluation of IBE. Examples of higher order designs are illustrated in Table 3.10 In higher order designs, each study subject receives each treatment formulation more than one time in a sequence or period hence such designs are called replicate crossover designs. The more sophisticated the design becomes, the more information for assessing bioequivalence can be extracted.

Table 3.10 Examples of 2×3 and 2×4 Crossover Designs

Three period crossover design

		Period		
		1	2	3
Sequence	1	**T**	**R**	**T**
	2	**R**	**T**	**R**

Four period crossover design

		Period			
		1	2	3	4
Sequence	1	**T**	**R**	**T**	**R**
	2	**R**	**T**	**R**	**T**

To estimate the within and between subject variances, components which are necessary when finding IBE, the appropriate design is the replicate design. Pharmaceutical companies and bioequivalence regulatory bodies around the globe prefer using the 2×2 crossover design for assessing ABE, however, this design is

not adequate to estimate the subject by formulation interaction (Jones and Kenward, 2003). IBE is discussed in this research report but cannot be evaluated since some necessary components for assessing IBE cannot be obtained from a standard 2×2 crossover design.

3.14 Effect Sizes

When presenting the results of a statistical research (or study), in addition to null hypothesis significance tests, researchers are encouraged to also report effect sizes (ES) as well as their corresponding confidence intervals. Kelley and Preacher (2012) is for the above view, that is, of including effect sizes and confidence intervals when reporting research results as emphasized by methodologists, journal editors, reviewers and professional organizations. Elsayir (2012) states that effect sizes have been available for decades, though not reported by statisticians when presenting results of their findings from research studies. He also indicated that the concept of effect sizes is actually from meta-analysis. Meta-analysis refers to a branch in social sciences and statistics where results or information from past research are used to motivate future studies (Nandy, 2012). Becker (2000) defines meta-analysis as a summary of previous research findings that uses quantitative methods to compare outcomes from different studies. Coe (2012) also indicated that effect sizes have been available for at least sixty years, though most of the literature on statistics does not cover effect sizes, with the exception of the books or material on meta-analysis. ES are rarely taught in introductory courses on statistics. Effect sizes are useful because:

- They enable researchers to report the magnitude of the difference of the means between two treatment effects by a unit less measure. It becomes easy for researchers to explain whether the results are practically viable or not basing their opinion on the actual size of the unit less ES, instead of only reporting the statistical significance;
- Effect sizes enable researchers to make comments on meta-analysis findings by comparing standard effect sizes from different bioequivalence studies;

- Effect sizes from past studies can be used in the planning stages of another study;
- Effect sizes help in sample size calculations for any study (Coe, 2012 and Nandy, 2012).

A statistical hypothesis is an assumption about the actual value of a population parameter, whereby there is the null hypothesis, H_0 (a claim) and the alternative hypothesis, H_α (used to ascertain as to whether the claim is valid or not). Hypothesis testing refers to the formal procedure used by statisticians or researchers in related fields to reject or fail to reject the null hypothesis (claim). Results which indicate that H_0 is reject (statistical significant) imply there is a difference, say, between means of two different groups (claim not valid). However, the magnitude or size of the difference is not known, hence, there is a need then to quantify the difference which leads to the calculation of effect sizes and their corresponding confidence intervals (Kelly and Preacher, 2012; Becker, 2000; Nandy, 2012 and Elsayir, 2012).

Researchers define effect sizes in a similar way however, there are variations in some of the definitions. The definition of effect size is split into the terms effect and size, then combined later as: Effect is a change or state of change caused by somebody or something while size refers to the degree of how huge or small a quantity is. Combining the two, effect size refers to expressing the difference of treatment means in terms of a specific value (Nandy, 2012). Effect size is a family of metrics that measure the magnitude of the difference between the treatment effects. ES differ from significance tests in the sense that they are independent of sample size (Becker, 2000 and Elsayir, 2012). Effect sizes are a family of indices that assign a number to the size of the difference between treatment effects and are used to address a question of interest (Kelly and Preacher, 2012).

The definitions of ES by Nandy (2012), Becker (2000) and Elsayir (2012) are basically the same, however, the one by Kelly and Preacher (2012) is broad since it links effect size with the question of interest. Effect sizes determine the practical importance of a study. Effect size analysis actually compares the mean of the test

treatment group with the mean of the reference treatment group. Statistical significance does not indicate the magnitude of the effect hence silent on the practical significance of a study (Steyn, 1999). Elsayir (2012) also talks of effect sizes being a useful method used when presenting findings from a research study and interpreting how the study is effective. ES has many advantages when compared to tests of statistical significance. They actually propel researchers to report results moving beyond the null hypothesis tests.

Confidence intervals give an indication as to whether two treatment means from different groups are the same or not, just like the hypotheses tests but the difference between an effect size and confidence interval is that an ES is an index that measures the magnitude of the difference of the means between treatments given as a range while confidence interval reflects the degree of confidence for having the magnitude of the difference of means in the interval. That is, effect size is a value, of which guidelines are then available, according to the various indices, on how to interpret them whereas given a confidence interval, one can indicate, with confidence, that a percentage of a particular statistic would lie in that interval (Steyn, 1999).

Effect sizes have the following facets, where facet relates to some characteristics of the effect size that relates to the manner by which the term is used. Facets of ES include:

- dimension,
- measure or index,
- value (Kelly and Preacher, 2012).

Dimension reflects to the type of information of interest so as to identify the appropriate units to be used. In real life there are a variety of dimensions, for example in Physics, dimensions include distance, weight, density, force and many more, of which all of them have different units. Variance is an example of an ES in statistics. Variance and other related measures of dispersion (spread) are unit less, hence effect size dimension gives a roadmap on how a research question will

be addressed in terms of the difference in magnitude (generally mean values) (Elsayir, 2012 and Kelly and Preacher, 2012).

Effect size measure, also known as effect size index, gives the formula that is suitable to calculate ES for a given situation (Lakens, 2013).

Effect size value is the number obtained when formulae has been used on bioequivalence data to obtain some statistics of interest. It is the magnitude of the effect size which is then used when reporting research findings (Becker, 2000).

Effect sizes are measured using a variety of formula depending on the situation (or study) since they vary from study to study. The two main ways of measuring effect sizes are:

- By calculating the difference between the means;
- By finding the relationship between the independent variable and the respective measurements of the dependent variable (Kelley and Preacher, 2012).

In essence, effect sizes are either reported as the magnitude of an effect or correlation coefficient, that is, the strength of the relationship between variables.

The necessary characteristics of ES are that:

- Computed numerical values need to be compared from different bioequivalence studies;
- Ability to calculate the standard error;
- Should not be directly linked to sample size (Nandy, 2012).

The fact that ES is not affected by sample size is also shared by Becker (2000) and Elsayir (2012).

Mathematically, the effect size is the magnitude of the difference between the actual of the statistic and the value specified in the null hypothesis. In other words, it is the difference between two means, generally, mean of the generic drug

minus the mean of the reference drug divided by the standard deviation of the two types of drugs. The division by the standard deviation is necessary since it enables researchers to compare effect sizes across experiments. It should be noted that means on their own simply gives the differences, on average, but does not say anything about the dispersion (spread) of the difference between the means. Types of effect sizes are discussed by researchers such as Nandy (2012), Becker (2000), Elsayir (2012), Warmbrod (2001), Lakens (2013) and Coe (2002).

3.14.1 Mean Differences Between Group Effect Sizes

3.14.1.1 Cohen's, *d*

Cohen's *d* is calculated by t-tests if given two independent samples. These effect sizes fall in the interval from $-\infty$ to ∞ and their interpretation is based on the number of standard deviations.

$$d = \frac{\bar{x}_T - \bar{x}_R}{S_{pooled}} \text{ where } S_{pooled} = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2}} \qquad (3.14.1.1.1)$$

3.14.1.2 Hedge's, *g*

Compute ES using Hedge's, *g* if the sample size is small. Hedge's compares the means of two groups where the mean differences have been standardized. Some authors on ES state that the Cohen's, *d* is also known as the Hedge's *g* effect sizes (Warmbrod, 2001; Coe, 2002 and Elsayir, 2012). While others such as Becker (2000), Nandy (2012) and Lakens (2013) view these two types effect sizes as different. The slight difference between the Cohen's, *d* and the Hedge's *g* effect sizes is the minus two on the denominator of the pooled standard deviation.

$$g = \frac{\bar{x}_T - \bar{x}_R}{S_{pooled}} \text{ where } S_{pooled} = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} \qquad (3.14.1.2.1)$$

3.14.1.3 Glass's, δ

Effect sizes are calculated by the Glass's, δ if the sample has unequal variances for given groups. The sample standard deviation of the reference group only is used so that effect size would be within accepted limits (Becker, 2000).

$$\delta = \frac{\bar{x}_T - \bar{x}_R}{S_R} \qquad (3.14.1.3.1)$$

General guidelines for interpretation of effect sizes:

- Less than or equal to 0.20 is a small effect size, where the variance explained is 1%;
- The magnitude of a moderate effect size is 0.50, with the variance explained being 10%;
- Greater than or equal to 0.80 is a large effect accounting for 25% variance explained (Cohen, 1992).

Cohen (1992) interpreted effect sizes as either being: small, medium and large depending on the assumption that both the reference and test groups are approximately normally distributed.

3.14.2 Correlation or Regression Effect Sizes

3.14.2.1 Pearson's Correlation Coefficient

Pearson's, *r*

$$r_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma_i^n (x_i - \bar{x})^2}\sqrt{\Sigma_i^n (y_i - \bar{y})^2}} \text{ where } -1 \ll r_{xy} \ll 1 \qquad (3.14.2.1.1)$$

Pearson's, *r* is ES is used to indicate the strength of the relationship between two variables.

Pearson's, R^2

Pearson's R^2 is known as the coefficient of determination gives the amount variation, in percentages, accounted for by the linear regression model. R^2 gives an indication as to how the model fits data (Nandy, 2012).

3.14.2.2 Cohen's, f^2

$$f^2 = \frac{R^2}{1-R^2} = \frac{\eta^2}{1-\eta^2} \qquad (3.14.2.2.1)$$

is used in multiple linear regression where $R^2 = \eta^2$. The standardized ES is the amount of variation explained by the model over the amount not explained. Cohen's f^2 is a biased estimate and tend to overestimate the ES for ANOVA. The unbiased estimate is called Omega-squared (Elsayir, 2012).

3.14.3 Contigency Tables Effect Sizes

Table 3.11 2×2 Table showing Smokers, Non-Smokers and their Disease Status

Risk	Disease Status	
	Present	Absent
Smokers	a	b
Non-smokers	c	d

3.14.3.1 Odds Ratio (OR)

$$OR = \frac{ad}{bc} \qquad (3.14.3.1.1)$$

Odds Ratio is used in situations where there are binary or categorical outcomes. OR values ranges from zero to infinity. If OR is greater than one, there is an increase in the odds relative to the reference group and a decrease in odds is attained if OR is less than one (Wilson, 2011).

3.14.3.2 Relative Risk (RR)

$$RR = \frac{a/(a+b)}{c/(c+d)} \qquad (3.14.3.2.1)$$

RR is a measure of risk relative to the independent variable. Relative risk is approximately equal to OR for small probabilities. Given as:

$$\frac{a/(a+b)}{c/(c+d)} \approx \frac{ad}{bc} \qquad (3.14.3.2.2)$$

The risk of disease X among smokers in Table 3.11 is equal to the relative risk times the risk of disease X among non-smokers if RR is greater than one (Wilson, 2011).

3.14.4 ANOVA or GLM Effect Sizes

3.14.4.1 Eta-Squared, η^2 and Partial Eta-Squared, η_p^2

Both are measures that estimate the association between variables of given samples.

$$\eta^2 = \frac{SS_{treatment}}{SS_{total}}; \qquad \eta_p^2 = \frac{SS_{treatment}}{SS_{total}+SS_{error}} \text{ where } 0 \ll \eta^2 \ll 1 \qquad (3.14.4.1.1)$$

these ES standardizes the amount of variance shared by the continuous and categorical outcomes. Partial Eta-Squared accounts for the percentage of the variance in the dependent variable explained by the variance in the independent variable. The interpretation of these effect sizes is synonymous to that of R^2 (coefficient of determination) in linear regression. Eta-squared is biased and generally overestimates the variance explained in the population. However, as the sample size increases, Eta-squared decrease (Lakens, 2013).

3.14.4.2 Omega-Squared, ω^2

$$\widehat{\omega}^2 = \frac{SS_{treatment} - df_{treatment} \times MS_{error}}{SS_{total} + MS_{error}} \quad (3.14.4.2.1)$$

Omega-squared estimate the amount of variance in the population that is explained by the treatment. Omega-squared is always smaller than η^2 or η_p^2 since Omega relates to the population and Eta measures the sample variance (Warmbrod, 2001).

3.14.4.3 Intraclass Correlation Coefficient

Intra correlation coefficient (ICC) is used for finding inter-rater reliability for two or more raters, though can also assess test-retest reliability. The ratio between group variance to the total variance is also a measure of ICC.

$$ICC = \frac{MS_{treatment} - MS_{error}}{MS_{treatment} + (n-1)MS_{error}} \quad (3.14.4.3.1)$$

ICC is interpreted in a similar way as Omega-squared (Warmbrod, 2001).

3.14.5 Chi-Square Tests Effect Sizes

3.14.5.1 Phi, Φ

$$\Phi = \sqrt{\frac{\chi^2}{n}} \quad (3.14.5.1.1)$$

used when there are crosstabs or chi-square tests specifically to test for the equality of proportions or tests of independence between two binary variables. Phi effect sizes are similar to correlation and Cohen's, *d* since they all measure the relationship between variables. The interpretation of the phi effect sizes is like that for Pearson's, *r* and R^2 (Nandy, 2012).

3.14.5.2 Cramer's, Φ or V

Cramer's phi can be used with categorical variables with more than two categories given contingency tables.

$$\Phi_c = \sqrt{\frac{\chi^2}{N(k-1)}} \qquad k=\min(R;C) \qquad (3.14.5.2.1)$$

measures the inter-correlation of the variables, but is biased since it increases with the number of cells. An increase in the number of rows and columns is an indication of a strong correlation between variables (Nandy, 2012).

Table 3.12 Magnitude of Effect Sizes (Nandy, 2012, pp 28)

Effect Size	**Small**	**Medium**	**Large**
r	0.10	0.30	0.50
r^2	0.01	0.09	0.25
η^2	0.01	0.06	0.14
R^2	0.01	0.06	0.14
Cohen's, *d*	0.20	0.50	0.80
Cramer's, *V*	0.10	0.30	0.50
Cohen's, f^2	0.02	0.15	0.30
OR	1.44	2.47	4.25

The following is relevant regarding the choice of effect size preferred in a given study:

- If all studies in the analysis are based on the same kind of data (means, binary or correlational), the researcher should select an effect size based on that kind of data;
- When some studies use means, other studies use binary data and some studies use correlational data, formulae to convert among effect sizes can then be applied;

- Studies that use different measures may differ from each other in substantive ways hence there is a need to consider this possibility when deciding if it makes sense to include the various studies in the same analysis (Kelly and Preacher, 2012).

When conversions are done to different ES measures, certain assumptions need to be made about the nature of the underlying effects. It should be noted that even if these assumptions do not hold exactly, the decision to use conversions is often better than the alternative, which is simply to omit the studies that happened to use a different metric since this would involve loss of information and possibly resulting in a biased sample of studies. Conversions were illustrated by Nandy (2012) moving from correlation, chi-square and odds ratio effect sizes to Cohen's, *d* effect sizes, the formulae are given in Table 3.13.

Table 3.13 Effect Size Conversions

Effect Size	Converted to Cohen's (d)
Correlation	$d = \frac{2r}{\sqrt{1-r2}}$
Chi-Square df=1 df > 1	$d = \sqrt{\frac{4x^2}{N-x^2}}$ $\sqrt{\frac{4x^2}{N}}$
Odds Ratio	$d = \frac{\text{In (0R)}}{1.81}$

It should be noted that while it is a noble idea to report effect sizes when presenting findings of a research study as suggested in this section, there are, however limitations:

- A problem may arise when interpreting standardized effect sizes if there is a limit on the sample size;

- Where the ES is not calculated using data which follows a normal distribution;
- When the ES is derived from a measure with an unknown reliability;
- There can also be issues related to variance estimate, notably if the methods of calculating effect sizes may have an influence on the estimate of the variance, for example, if the difference between the means is computed from dichotomous data or if the numerator of the differences is adjusted for baseline or other covariates or the study data involves clusters (Coe, 2012).

ES were calculated in this study on results found on ABE by using the C.I. approach, TOST, Wilcoxon Mann-Whitney test statistics and PBE. ES were also calculated on results obtained after bootstrapping.

Results obtained after assessing ABE and PBE between the test and reference treatment formulations are outlined and analysed in the next chapter of this study. Outliers were checked and the sample size required to achieve the required power is also discussed. Effect sizes were used to quantify the difference between the ratios of means of the two treatment formulations.

Chapter 4

Results and Analysis

4.1 Introduction

Pharmacokinetic parameters T_{max}, C_{max}, AUC_{0-t} and $AUC_{0-\infty}$ obtained from sheep 1 to 14 are in Tables 4.1, 4.2, 4.3 and 4.4. Steps used for calculating $AUC_{0-\infty}$ are in Appendices 1, 2, 3, 4 and 5. Data sets in tables mentioned above were used to compare the bioavailability between the test treatment and reference treatment formulations in this research report.

Outliers are identified using the Cook's likelihood distance test, though will not be deleted from the data set as recommended by the FDA (1992) and Patterson (2010). If the power of a test is at least 80%, the bioequivalence study can be done efficiently, saving on time and resources since the appropriate number of study subjects will be found prior to the commencement of the study. Section 4.3 contains SAS 9.4 outputs illustrating the effect of sample size in determining the power of a test. By achieving the relevant sample size to obtain desired power of a test, of at least 80%, pharmacokinetic parameters for the test treatment and reference treatment formulations are then tested for bioequivalence. ABE is assessed through the classic confidence interval approach and the interval hypothesis testing approach using Schuirmann's TOST procedures. Results for the parameter T_{max} are obtained using the nonparametric test, Wilcoxon Mann-Whitney two one-sided test statistics. In this chapter, the test treatment and reference treatment formulations are shown to be population bioequivalent (PBE) by using the relevant criteria. Bootstrapping, a resampling method which increases the sample size is done and pharmacokinetic parameters $AUC_{0-\infty}$ and C_{max} are shown to be average bioequivalent by the classic confidence interval approach. Effect sizes are covered in Section 4.7.

Table 4.1 Values of T_{max} and C_{max} in Sequence 1

	Animal	T_{max}	C_{max}	lnC_{max}
Test	5	2	7.3284	1.991757
	6	2	9.4104	2.241815
	9	9	21	3.044522
	10	2.02	13.5789	2.608517
	11	2	9.5904	2.260763
	12	1	8.4611	2.135479
	14	1	11.4642	2.439229
Reference	5	2	11.9468	2.480463
	6	1	16.2239	2.786485
	9	2	10.3462	2.336619
	10	2	12.8158	2.550679
	11	1	8.0715	2.088339
	12	0.5	14.7725	2.692767
	14	0.5	15.699	2.753597

Table 4.2 Values of T_{max} and C_{max} in Sequence 2

	Animal	T_{max}	C_{max}	lnC_{max}
Reference	1	2	17.303	2.85088
	2	2	8.5327	2.143906
	3	2	11.0619	2.403507
	4	2	10.7164	2.371775
	7	2	9.7471	2.27697
	8	2	11.4669	2.439465
	13	1	10.2324	2.325559
Test	1	2	11.5974	2.450781
	2	2	9.3668	2.237172
	3	4	11.1924	2.415235
	4	2	12.6993	2.541547
	7	2	11.2346	2.418998
	8	2	12.7964	2.549164
	13	2	12.8171	2.55078

The crossover design, Figure 3.2, was used at the University of Pretoria at Onderstepoort within the Directorate of Veterinary Pharmacology where randomization was done such that some sheep ended up in sequence 1 while others in sequence 2. Table 4.1 shows values of T_{max} and C_{max} from the

bioequivalence data obtained, courtesy of the University of Pretoria after dosing sheep 5, 6, 9, 10, 11, 12 and 14 in sequence 1 with the test treatment formulation in period 1, an adequate washout period was allowed to pass, then sheep were dosed with the reference treatment formulation in period 2. Table 4.2 indicates values of T_{max} and C_{max} for sheep 1, 2, 3, 4, 7, 8 and 13 in sequence 2 after dosing with the reference treatment formulation in period 1, a long enough washout period was taken then sheep were dosed with the test treatment formulation in period 2.

Table 4.3 Calculating $AUC_{0-\infty}$ for the sheep in Sequence 1

Sequence 1

	Animal	AUC_{0-t}	C_t	λ	$AUC_{0-\infty}$	$lnAUC_{0-\infty}$
Test	5	132.0207	0.0195	0.192945	132.1218	4.883724
	6	170.1138	0.2103	0.130879	171.7206	5.145869
	9	222.0024	0.209	0.07494	224.7913	5.415172
	10	173.6732	0.0075	0.222355	173.7069	5.15737
	11	119.2981	0.1943	0.084889	121.587	4.80063
	12	138.9857	0.0972	0.095022	140.0086	4.941704
	14	190.5718	0.0664	0.160888	190.9845	5.252192
Reference	5	170.5818	0.0597	0.10755	171.1369	5.142464
	6	196.4537	0.0066	0.242045	196.481	5.280566
	9	207.0228	0.1126	0.145411	207.7972	5.336562
	10	159.9659	0.0845	0.179335	160.4371	5.077902
	11	119.1357	0.0376	0.162292	119.3674	4.782206
	12	153.0361	0.1695	0.154854	154.1307	5.037801
	14	220.8329	0.1949	0.151353	222.1206	5.403221

AUC was calculated using formulae 3.5.1 and 3.5.2, the procedure for the calculations are outlined in Appendices 1, 2, 3, 4 and 5. Table 4.3 shows values of $AUC_{0-\infty}$ after sheep 5, 6, 9, 10, 11, 12 and 14 were dosed with the test treatment formulation in period 1 followed by the reference treatment formulation in period 2. Table 4.4 is for values of $AUC_{0-\infty}$ for sheep 1, 2, 3, 4, 7, 8 and 13 in sequence 2 dosed with the reference treatment formulation in period 1 then the test treatment formulation in period 2.

Table 4.4 Calculating $AUC_{0-\infty}$ for the sheep in Sequence 2

Sequence 2

	Animal	AUC_{0-t}	C_t	λ	$AUC_{0-\infty}$	$lnAUC_{0-\infty}$
Reference1	1	167.1385	0.0336	0.19534	167.3105	5.119851
	2	166.5281	0.1206	0.146724	167.3501	5.120088
	3	181.6076	0.0852	0.166944	182.118	5.204655
	4	152.5701	0.1561	0.148981	153.6179	5.034468
	7	174.1533	0.0079	0.256715	174.1841	5.160113
	8	152.1165	0.0507	0.192024	152.3805	5.026381
	13	195.1567	0.0985	0.117522	195.9948	5.278088
Test	1	170.3735	0.1173	0.162799	171.094	5.142213
	2	164.6649	0.0885	0.119894	165.4031	5.108385
	3	187.0004	0.1407	0.15338	187.9177	5.236004
	4	184.3159	0.0517	0.192347	184.5847	5.218108
	7	201.7071	0.1005	0.123786	202.519	5.310834
	8	183.1658	0.0099	0.22369	183.2101	5.210633
	13	214.0738	0.0992	0.172126	214.6501	5.369009

4.2 Outliers

4.2.1 Outliers for *AUC* from time zero to infinity

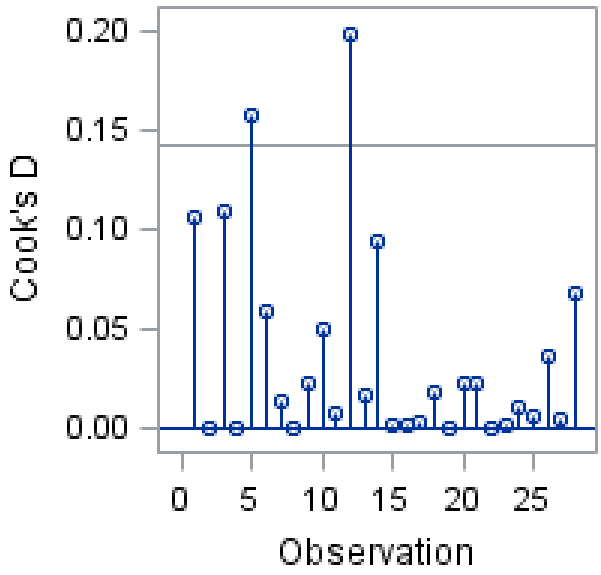


Figure 4.1 Outliers for the Parameter $AUC_{0-\infty}$

Observations 5 and 12 are outliers for the pharmacokinetic parameter $AUC_{0-\infty}$, transformed data, shown in Figure 4.1. Considering $AUC_{0-\infty}$ for the untransformed data, Appendix 7, Figure 2 there are also two outliers. The same number of outliers for the transformed and untransformed data identified could be attributed to the fact that untransformed data was nearly symmetric, hence transforming did not change the data structure much. The histogram in Appendix 7, Figure 1 is almost symmetric, confirming that the original data was almost symmetric.

Observation 3 is the only outlier for the pharmacokinetic parameter C_{max} detected in Figure 4.2. Comparing the outlier outputs for the transformed and untransformed data, Figure 2, Appendix 8 exhibits three outlying observations whereas for the transformed data there is only one outlier. The histogram in Figure 1, Appendix 8 is skewed to the right which explains the need to transform the data and the difference of the number of outliers identified using transformed and original data.

4.2.2 Outliers for C_{max}

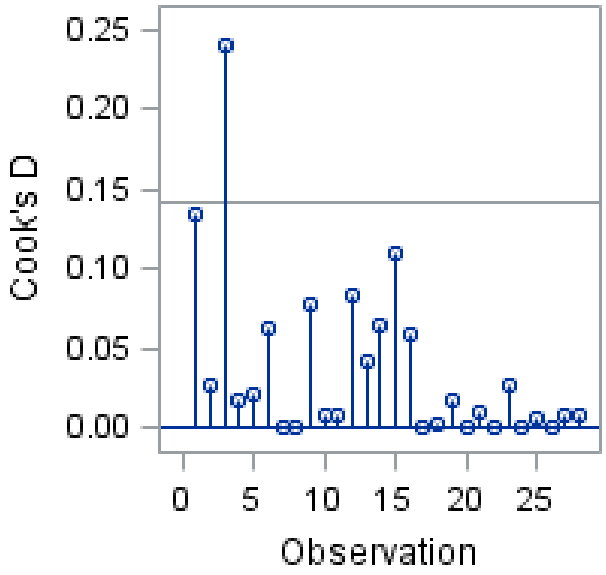


Figure 4.2 Outliers for the Parameter C_{max}

The Cook's likelihood distance test was used in this research report because it clearly identifies the outliers, indicating the exact number of the observation which is an outlier graphically. Suppose other methods for detecting outliers were used, ED, HT or MS, the same outliers could have been identified for the parameters $AUC_{0-\infty}$ and C_{max} though in a different format, not as figures obtained using the LD.

Outliers can be detected from a dataset but cannot be removed for the evaluation of bioequivalence (FDA, 1992 and Patterson, 2010). An outlier detection test can be done but this is not necessarily an expectation of the medicine agencies. Deleting an outlier is unacceptable since outliers do not usually affect the overall results on bioequivalence because the individual observations form a negligible portion of the overall average results, for example, when calculating the parameter *AUC* the contribution of say one or two observations does not affect the *AUC* value that much. The inference on bioequivalence done by various regulatory authorities is based on the complete data set. The implication for not removing outlying observations when doing data analysis and evaluation is that in practice outliers do not exist in a bioequivalence data set. It is recommended that statisticians need to check for the normality assumptions in their bioequivalence models, though, they are advised not to spend too much time since outliers do not contribute much on bioequivalence results. The opinion of other authors on bioequivalence is that the decision of deleting outlying observations in a data set depends on the discretion of different bioequivalence regulatory bodies (Patterson, 2010).

4.3 Power

Table 4.5 Power $AUC_{0-\infty}$

Mu1	Mu2	StDev1	StDev2	Corr	N
5.1432	5.1566	0.04412	0.04412	0.69172	28
					42
					56

Alpha	StDevDiff	NCP	Critical Value	Power
0.05	0.034644	4.18906	4.21001	0.50552
		6.28360	4.07855	0.68712
		8.37813	4.01620	0.81156

Table 4.6 Power C_{max}

Mu1	Mu2	StDev1	StDev2	Corr	N
2.4644	2.4204	0.0665	0.0665	-0.11476	28
					42
					56

Alpha	StDevDiff	NCP	Critical Value	Power
0.05	0.099295	5.49805	4.21001	0.61826
		8.24708	4.07855	0.80065
		10.9961	4.01620	0.90279

In Tables 4.5 and 4.6, Mu1 and Mu2 are the least squares mean for the test and reference treatments respectively while StDev1 and StDev2 are the respective standard deviations for the test and reference treatment formulations. For the parameter *AUC* the power of the test increases from 0.50552 to 0.68712 and eventually to 0.81156 shown in Table 4.5 as the sample size increases from 28, 42

and 56 respectively. Considering the parameter C_{max} as the sample sizes increase from 28 to 42 and to 56, the respective power increases from 0.61826 to 0.80065 and finally to 0.90279 in Table 4.6.

This confirms the discussion in Section 3.9 stating that as the sample size increases, the power of a test also increases. This strategy of increasing the power is used if the power is low. Increasing the power of a test by taking a larger sample size improves the ability to find the difference in means of the treatment formulations when they indeed exist. As the population variance gets smaller, the power of a test improves by increasing and vice versa and as differences in means increases the power increases, but the researcher has no control over the population variances or difference in means. The feasible way for the researcher to improve the power of a test if it is low, is to increase the sample size (Elsayir, 2012).

4.4 Average Bioequivalence

4.4.1 Confidence Intervals Approach

4.4.1.1 Classic Confidence Interval of the Difference

The ANOVA Table 4.7 is used to find the variance (mean square error) which is half of 0.00565711 obtained from model 3.4.1 and needed when calculating confidence intervals (C.I.) using formula 2.7.2.1.1. The variance for a 2×2 crossover design is multiplied by 0.5, however, if a parallel design is used the MSE in Table 4.7 is captured as it is for purposes of calculating C.I. (Concordet, 2004). Multiplying the MSE by 0.5 is done in all cases where the ANOVA using GLM procedure is done, that is, in Tables 4.10 and 4.13.

Table 4.7 ANOVA for AUC_{0-t} using the GLM Procedure in SAS 9.4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	0.66681567	0.04445438	7.86	0.0005
Error	12	0.06788537	0.00565711		
Corrected Total	27	0.73470104			

R-Square	Coeff Var	Root MSE	$lnAUC_{mean}$
0.907601	1.461718	0.075214	5.145576

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	1	0.00092452	0.00092452	0.16	0.6931
Period	1	0.04656188	0.04656188	8.23	0.0141
Seq	1	0.02983240	0.02983240	5.27	0.0405
Animal	12	0.58949687	0.04912474	8.68	0.0004

Table 4.8 ANOVA for AUC_{0-t} using Mixed Procedure in SAS 9.4

Least Squares Means								
Treatment	Estimate	S.E.	DF	t Value	Pr > |t|	Alpha	Lower Bound	Upper Bound
R	5.1398	0.04423	12	116.20	<.0001	0.1	5.0610	5.2187
T	5.1513	0.04423	12	116.46	<.0001	0.1	5.0725	5.2302

ANOVA Table 4.8 provides the mean for the reference treatment, 5.1398 and the mean for the test treatment, 5.1513.

Table 4.9 Confidence Intervals (C.I.) for *AUC* from time zero to time *t*

	Standard error	Lower Bound	Lower 90% Confidence Limit	Upper 90% Confidence Limit	Upper Bound
C.I. for difference of means	0.02843	−0.2231	−0.03917	0.06217	0.2231
C.I. for ratio of means		0.8	0.96159	1.06414	1.25

Average bioequivalence of the two treatment drug formulations has been achieved at 5% significance level using classic C.I. of the difference of the means as well as the classic C.I. of the ratio of the means since for the C.I. of the differences, the limits −0.03917 and 0.06217 in Table 4.9 lie entirely in the acceptance range, −0.2231 to 0.2231. For the C.I. on the ratios, 0.96159 and 1.06414 in Table 4.9 lie entirely in the acceptance interval of 0.8 to 1.25.

Therefore the test treatment and reference treatment formulations are average bioequivalent using the pharmacokinetic parameter *AUC* from time zero to time *t* by the classic confidence interval approach.

Table 4.10 ANOVA for $AUC_{0-\infty}$ using GLM Procedure in SAS 9.4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	0.65811748	0.04387450	7.58	0.0006
Error	12	0.06941768	0.00578481		
Corrected Total	27	0.72753516			

R-Square	Coeff Var	Root MSE	$lnAUC_{mean}$
0.904585	1.476892	0.076058	5.149865

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	1	0.00125533	0.00125533	0.22	0.6497
Period	1	0.04444893	0.04444893	7.68	0.0169
Seq	1	0.02774817	0.02774817	4.80	0.0490
Animal	12	0.58466505	0.04872209	8.42	0.0004

The MSE for the parameter *AUC* to infinity is half of 0.00578481 in Table 4.10 obtained using model 3.4.1.

Table 4.11 ANOVA for $AUC_{0-\infty}$ using Mixed Procedure in SAS 9.4

Least Squares Means								
Treatment	Estimate	S.E.	DF	t Value	Pr > |t|	Alpha	L.B.	U.B.
R	5.1432	0.04412	12	116.57	<.0001	0.1	5.0645	5.2218
T	5.1566	0.04412	12	116.87	<.0001	0.1	5.0779	5.2352

The mean for the reference treatment drug is 5.1432 while the one for the test treatment formulation is 5.1566, both in Table 4.11.

Table 4.12 Confidence Intervals for *AUC* from time zero to infinity

	Standard error	Lower Bound	Lower 90% Confidence Limit	Upper 90% Confidence Limit	Upper Bound
C.I. for difference of means	0.02875	−0.2231	−0.03784	0.06464	0.2231
C.I. for ratio of means		0.8	0.96287	1.06677	1.25

The lower confidence limits and upper confidence limits for the mean differences of the parameter *AUC* from time zero to infinity, −0.03784 and 0.06464 in Table 4.12 for the two treatment formulations are totally included within the average bioequivalence bounds −0.2231 and 0.2231.

Similarly, for the confidence intervals for the ratio of the means, 0.96287 and 1.06677 in Table 4.12 lies entirely within the average bioequivalence range, 0.8 to 1. 25.

Average bioequivalence of the two treatment formulations, test and reference formulations, is therefore concluded for the parameter *AUC* from time zero to infinity using the classic confidence interval approach.

Table 4.13 ANOVA for C_{max} using GLM Procedure in SAS 9.4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	0.76395940	0.05093063	0.76	0.6965
Error	12	0.80427853	0.06702321		
Corrected Total	27	1.56823793			

R-Square	Coeff Var	Root MSE	$lnC_{maxmean}$
0.487145	10.59982	0.258888	2.442385

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	1	0.01351912	0.01351912	0.20	0.6613
Period	1	0.06208561	0.06208561	0.93	0.3548
Seq	1	0.00676723	0.00676723	0.10	0.7561
Animal	12	0.68158744	0.05679895	0.85	0.6105

The MSE for the parameter C_{max} is half of 0.06702321 in Table 4.13 found using the model 3.4.1.

Table 4.14 ANOVA for C_{max} using Mixed Procedure in SAS 9.4

Least Squares Means								
Treatment	Estimate	S.E.	DF	t Value	Pr > |t|	Alpha	L.B.	U.B.
R	2.4644	0.06650	12	37.06	<.0001	0.1	2.3458	2.5829
T	2.4204	0.06650	12	36.40	<.0001	0.1	2.3019	2.5389

From Table 4.14, the least mean squares 2.4644 and 2.4204 for the reference treatment and test treatment, respectively, are obtained.

Table 4.15 Confidence Intervals for C_{max}

	Standard error	Lower Bound	Lower 90% Confidence Limit	Upper 90% Confidence Limit	Upper Bound
C.I. for difference of means	0.09785	−0.2231	−0.21860	0.13040	0.2231
C.I. for ratio of means		0.8	0.80364	1.13928	1.25

Average bioequivalence of the two treatments has been achieved using the classic confidence interval approach for the difference of means of parameter C_{max} because both confidence limits −0.21840 and 0.13040 in Table 4.15 are totally included in the acceptance range of −0.2231 to 0.2231.

Considering the C.I. for the ratio of means, average bioequivalence is also concluded since the confidence limits 0.80364 and 1.13928, Table 4.15, lies entirely in the acceptance range of 0.8 to 1.25.

The test treatment and reference treatment formulations have been shown to be average bioequivalent using the classic confidence interval approach.

4.4.2 Interval Hypothesis Testing Approach

4.4.2.1 Schuirmann's Two One-Sided tests procedure

The generally accepted method of testing for ABE according to Jones and Kenward (2003) and Chow and Liu (2000) is the TOST procedure proposed by Schuirmann (1987).

The Schuirmann's two one-sided tests lower and upper limits, 0.9908 and 1.0153, calculated by using 2.7.3.1.4 are obtained from Table 4.16. The geometric mean is

applied hence the confidence limits are for the ratio of means. The lower and upper limits (above) lie entirely in the confidence acceptance range 0.8 to 1.25, which implies that average bioequivalence is achieved.

Using the hypotheses approach, average bioequivalence is achieved if H_0 is rejected on both sets, in favour of the alternative hypothesis. C.I. limits in Table 4.16 indicate that the null hypothesis of 2.7.3.1.2 is rejected on both sides of the test, therefore, average bioequivalence is concluded using the Schuirmann's two one-sided tests procedure at 5% significance level for the parameter *AUC* from time zero to infinity.

Table 4.16 TOST for $AUC_{0-\infty}$

N	Geometric Mean	Coefficient of Variation	Minimum	Maximum
14	1.0030	0.0258	0.9491	1.0363

Geometric Mean	95% CL Mean		Coefficient of Variation		95% CL CV	
1.0030	0.9881	1.0180	0.0258	0.0187	0.0415	1.0030

Geometric Mean	Lower Bound		90% CL Mean			Upper Bound	Assessment
1.0030	0.8	<	0.9908	1.0153	<	1.25	Equivalent

Test	Null	DF	t Value	P-Value
Lower	0.8	13	32.82	<.0001
Upper	1.25	13	-31.96	<.0001
Overall				<.0001

Table 4.17 TOST for C_{max}

N	Geometric Mean	Coefficient of Variation	Minimum	Maximum
14	0.9815	0.1474	0.7930	1.3030

Geometric Mean	95% CL Mean		Coefficient of Variation	95% CL CV	
0.9815	0.9018	1.0682	0.1474	0.1066	0.2396

Geometric Mean	Lower Bound		90% CL Mean			Upper Bound	Assessment
0.9815	0.8	<	0.9157	1.0520	<	1.25	Equivalent

Test	Null	DF	t Value	P-Value
Lower	0.8	13	5.22	<.0001
Upper	1.25	13	-6.17	<.0001
Overall				<.0001

Average bioequivalence is concluded using C_{max} as in Table 4.17, 0.9157 to 1.052 is totally included in the average bioequivalence interval, 0.8 to 1.25.

Considering hypotheses 2.7.3.1.2, both sets rejects the null hypothesis, which implies that average bioequivalence is found between the two treatment formulations at 5% level of significance using the Schuirmann's two one-sided tests procedures.

4.4.3 Wilcoxon Mann-Whitney Two One-Sided test statistics

For the pharmacokinetic parameter T_{max} the Wilcoxon Mann-Whitney test statistics are shown in Appendix 6 and the ±20 rule is used since the data is untransformed. T_{max} is sampled from discrete values, thus a nonparametric method is applied to assess ABE. R_L calculated using equation 2.7.4.10 and R_U are the sum of ranks of the response for sheep in sequence 1 and sum of ranks for

sheep in sequence 2 respectively. W_L and W_U the Wilcoxon Mann-Whitney test statistics for the null and alternative hypothesis respectively are calculated using equations 2.7.4.11 and 2.7.4.13 respectively.

Table 4.18 Wilcoxon Mann-Whitney Two One-Sided test statistics

R_L	77	R_U	57
W_L	49	W_U	29
$W_{0.95}$	37	$W_{7;7;0.05}$	40

Recall the hypothesis is as follows:

$H_{01}: \theta_L^* \leq 0$ vs. $H_{a1}: \theta_L^* > 0$ and $H_{02}: \theta_U^* \geq 0$ vs. $H_{a2}: \theta_U^* < 0$

Of which the decision rule says:

H_{01} is rejected if $W_L > w(1-\alpha)$ and we reject H_{02} if $W_U < w(\alpha)$, whereby average bioequivalence is achieved if both sets of hypotheses are rejected at 5% significance level.

The Wilcoxon Mann-Whitney two one-sided test statistics indicate that indeed H_{01} is rejected at 5% level of significance and H_{02} is also rejected. Average bioequivalence of the two treatments, the test and reference drug formulations is concluded since W_L=49 is greater than $w(1-\alpha)$=37 and W_U=29 is less than $w(\alpha)$=40.

4.5 Population Bioequivalence

Kenward and Jones (2003) suggested that PBE can be calculated using data either from the 2×2 crossover design or from a replicate design, as long as an appropriate mixed model is fitted. The same view is shared by the FDA (1999, 2001), whereby they state that the standard 2× 2 crossover design may be used for PBE.

ANOVA Tables 4.19, 4.20 and 4.21 on covariance parameter estimates AUC_{0-t}, $AUC_{0-\infty}$and C_{max} indicates FA(1,1) and FA(2,2) the total variance for the reference treatment formulation and total variance for the total variance of the test treatment formulation respectively. Least square means, reference and test treatment formulations for AUC_{0-t}, $AUC_{0-\infty}$ and C_{max} are obtained from Tables 4.8, 4.11 and 4.14 respectively. The values of FA(1,1) and FA(2,2) are substituted in equation 2.7.5.1, the reference-scaled moment based criterion for finding the population bioequivalence value to be compared to θ_p=1.7448, the population bioequivalence limit FDA (2001). The reference-scaled criterion is applied in all the three cases since $\sigma_R^2 > \sigma_O^2$, that is, the total variance for the reference drug formulation is greater than 0.04, the pre-specified constant total variance. If the specified constant total variance was less than the total variance for the reference treatment drug then the constant-scaled criterion 2.7.5.2 could have been used.

Table 4.19 ANOVA for Covariance Parameter Estimates for *AUC* from time zero to time *t*

Covariance Parameter Estimates			
Cov Parm	Subject	Group	Estimate
FA(1,1)	Animal		0.1610
FA(2,1)	Animal		0.1350
FA(2,2)	Animal		0.09814
Residual	Animal	Treatment R	0.000486
Residual	Animal	Treatment T	0.000518

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Seq	1	12	0.61	0.4509
Period	1	12	8.23	0.0141
Treatment	1	12	0.16	0.6931

Estimates								
Label	Estimate	S.E.	DF	t value	Pr > |t|	Alpha	L.B.	U.B.
test-ref	0.0114	0.0284	12	0.40	0.6931	0.1	-0.0391	0.0621

Table 4.20 ANOVA for Covariance Parameter Estimates for *AUC* from time zero to infinity

Covariance Parameter Estimates			
Cov Parm	Subject	Group	Estimate
FA(1,1)	Animal		0.1612
FA(2,1)	Animal		0.1311
FA(2,2)	Animal		0.1001
Residual	Animal	Treatment R	0.000530
Residual	Animal	Treatment T	0.000537

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Seq	1	12	0.52	0.4837
Period	1	12	6.97	0.0216
Treatment	1	12	0.29	0.5991

Estimates								
Label	Estimate	S.E	DF	t value	Pr > |t|	Alpha	L.B.	U.B.
test-ref	0.0158	0.0292	12	0.54	0.5991	0.1	-0.0363	0.0679

Table 4.21 ANOVA for Covariance Parameter Estimates for C_{max}

Covariance Parameter Estimates			
Cov Parm	Subject	Group	Estimate
FA(1,1)	Animal		0.2312
FA(2,1)	Animal		-0.02211
FA(2,2)	Animal		0.2594
Residual	Animal	Treatment R	0.002344
Residual	Animal	Treatment T	0.000233

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Seq	1	12	0.12	0.7359
Period	1	12	0.93	0.3548
Treatment	1	12	0.20	0.6613

Estimates								
Label	Estimate	S.E.	DF	t value	Pr > |t|	Alpha	L.B.	U.B.
test-ref	-0.0439	0.0978	12	-0.45	0.6613	0.1	-0.2183	0.1305

Table 4.22 compares population bioequivalence values (θ_{PBE}) for the pharmacokinetic parameters AUC_{o-t} , $AUC_{o-\infty}$ and C_{max}. All the θ_{PBE} values are less than 1.7448 (population bioequivalent limit) FDA (2001). Therefore population bioequivalence of the two treatment formulations, reference and test, has been achieved using the moment-based criterion 2.7.5.1.

Table 4.22 Evaluation of Population Bioequivalence

	θ_{PBE}	θ_P	Results	Conclusion
AUC_{o-t}	0.904515	1.7448	Reject H_o	Population Bioequivalent
$AUC_{o-\infty}$	0.906715		Reject H_o	Population Bioequivalent
C_{max}	1.070594		Reject H_o	Population Bioequivalent

The hypotheses criterion for assessing population bioequivalence equivalent to criterion 2.7.5.1 and 2.7.5.2 is:

H_0^{PBE}:$\theta_{PBE} \gg \theta_P$ versus H_1^{PBE}:$\theta_{PBE} < \theta_P$

Population bioequivalence is concluded if H_0^{PBE} is rejected.

In Table 4.22, the null hypothesis is rejected in favour of the alternative hypothesis therefore population bioequivalence is concluded for AUC_{0-t}, $AUC_{0-\infty}$ and for C_{max}.

4.6 Bootstrapping

Bootstrapping resamples are on Appendix 9. These bootstrap samples were obtained using the INDEX function on Excel. The function is: INDEX((range of cells), ROWS(range of cells)*RAND()+1, COLUMNS(range of cells)*RAND()+1) (Rochowicz, 2011).

4.6.1 Bootstrapping $AUC_{0-\infty}$

Table 4.23 ANOVA for the Bootstrap Samples of $AUC_{0-\infty}$ using the GLM Procedure in SAS 9.4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	0.13972056	0.00931470	0.42	0.9701
Error	96	2.13268671	0.02221549		
Corrected Total	111	2.27240727			

R-Square	Coeff Var	Root MSE	$lnAUC_{mean}$
0.061486	2.887142	0.149049	5.162497

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	1	0.00020412	0.00020412	0.01	0.9238
Period	1	0.02811500	0.02811500	1.27	0.2634
Seq	1	0.02184208	0.02184208	0.98	0.3239
Animal	12	0.08955937	0.00746328	0.34	0.9805

Table 4.24 ANOVA for the Bootstrap Samples of $AUC_{0-\infty}$ using the Mixed Procedure in SAS 9.4

Least Squares Means								
Treatment	Estimate	S.E.	DF	t value	Pr > |t|	Alpha	L.B.	U.B.
R	5.1638	0.01917	96	269.39	<.0001	0.1	5.1320	5.1957
T	5.1611	0.01917	96	269.25	<.0001	0.1	5.1293	5.1930

The MSE is half of 0.02221549 in Table 4.23 the least mean squares for the test and reference formulation are 5.1611 and 5.1638 respectively from Table 4.24. The calculated C.I for bootstrap samples is for the ratio of means. When calculating the bootstrap C.I. the critical values used are from the Normal distribution since the sample size is large, n=112. For a large sample the t-distribution approximates the Normal distribution. Table 4.25 indicates that the bootstrap C.I. for *AUC* from time zero to infinity is 0.98281 to 1.02119 which lies entirely in the confidence acceptance region 0.8 to 1.25. ABE is therefore concluded for the bootstrap samples by using the classic C.I. approach.

Table 4.25 Bootstrap C.I. for $AUC_{0-\infty}$

	Standard error	Lower Bound	Lower 90% Confidence Limit	Upper 90% Confidence Limit	Upper Bound
C.I. for ratio of means	0.02817	0.8	0.98281	1.02119	1.25

4.6.2 Bootstrapping C_{max}

Table 4.26 ANOVA for the Bootstrap Samples of C_{max} using the GLM Procedure in SAS 9.4

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	15	1.11089103	0.07405940	1.49	0.1228
Error	96	4.75928937	0.04957593		
Corrected Total	111	5.87018040			

R-Square	Coeff Var	Root MSE	$lnC_{maxmean}$
0.189243	9.183376	0.222657	2.424561

Source	DF	Type I SS	Mean Square	F Value	Pr > F
Treatment	1	0.09528051	0.09528051	1.92	0.1689
Period	1	0.00001635	0.00001635	0.00	0.9855
Seq	1	0.59020016	0.59020016	11.90	0.0008
Animal	12	0.42539402	0.03544950	0.72	0.7334

The MSE is 0.5×0.04957593 obtained in Table 4.26 while the least squares mean for the test treatment is 2.4537 and that of the reference treatment is 2.3954 in Table 4.27. The calculated C.I. for the ratio of means for the pharmacokinetic parameter C_{max} is 1.0371 to 1.09818 in Table 4.28. This confidence interval is totally included in the ABE acceptance region 0.8 to 1.25. Bootstrap samples of C_{max} have been proved to be average bioequivalent by the classic C.I. approach.

Table 4.27 ANOVA for the Bootstrap Samples of C_{max} the Mixed Procedure in SAS 9.4

Least Squares Means								
Treatment	Estimate	S.E.	DF	t value	Pr > |t|	Alpha	L.B.	U.B.
R	2.3954	0.02928	96	81.81	<.0001	0.1	2.3468	2.4440
T	2.4537	0.02928	96	83.81	<.0001	0.1	2.4051	2.5024

Table 4.28 Bootstrap C.I. for C_{max}

	Standard error	Lower Bound	Lower 90% Confidence Limit	Upper 90% Confidence Limit	Upper Bound
C.I. for ratio of means	0.04208	0.8	1.03710	1.09818	1.25

4.7 Effect Sizes

Table 4.29 Effect Sizes Values for Different Types of ES

Effect size	$AUC_{0-\infty}$	C_{max}
Cohen's *d*	2.316	0.656
Pearson's, *r*	0.692	0.115
Eta-Squared, η^2	0.002	0.018

Pearson's, *r* and Eta-Squared, η^2 ES are calculated using equations 3.14.2.1.1 and 3.14.4.1.1 of which, according to Cohen (1992), both ES are interpreted as small. This implies that there is an insignificant difference of bioavailability between the two treatment formulations. Eta's squared for $AUC_{0-\infty}$ effect size, 0.002, is also a small ES. However, the parameter exhibits a $AUC_{0-\infty}$ medium ES, 0.692, using Pearson's and a large ES, 2.316, using Cohen's, *d*. C_{max} also has a medium ES when calculated using Cohen's, *d*. Ideally, all ES should be small since for two treatment formulations to be bioequivalent, the magnitude of the difference of the means should be minimal.

A summary of the results discussed in this chapter follows in the next chapter. We consider all statistical tests carried out and then conclude as whether the aims and objectives of the study were achieved or not. Limitations of the study and possible improvements which can be done are also covered in the next chapter.

Chapter 5

Conclusion on Results and Recommendations

5.1 Introduction

The aims and objectives of this research report have been achieved since it has been established by statistical techniques that the test and reference treatment formulations are ABE and PBE. Generic drugs (test), which cost less as compared to the brand name drugs (reference) can be taken without much fear because it can be shown that two treatment formulations are bioequivalent hence reducing the cost of healthcare. The model used, 2×2 crossover design, is however inadequate for assessing IBE which gives room for the use of a higher order model.

5.2 Summary

Average bioequivalence between the test and reference drugs used in this research report is achieved for the pharmacokinetic parameters *AUC* and C_{max} by using the classical (shortest) confidence intervals, Schuirmann's two one-sided tests and interval hypotheses. ABE is also concluded from the bootstrap samples of $AUC_{0-\infty}$ and C_{max} by using the classic C.I. approach. For the parameter T_{max} evaluated by the nonparametric test, Wilcoxon Mann-Whitney two one-sided test statistics, we also conclusively arrive at the decision of average bioequivalence between the generic and brand name drug since both sides of the test suggests that the formulations are bioequivalent. Using the Wilcoxon Mann-Whitney two one-sided test statistics, both sides of the test should be average bioequivalent but if one is not, then bioequivalence is not wholly achieved.

Pharmacokinetic parameters, AUC_{0-t}, $AUC_{0-\infty}$ and C_{max} proved to be population bioequivalent because θ_{PBE} is less than 1.7448 (population bioequivalence limit). Using hypothesis 2.7.5.3, H_0^{PBE} is rejected, confirming that the two treatment formulations are bioequivalent. Outliers could not have had too much effect on the results as discussed in the analysis Section 4.2. Individual bioequivalence

could not be assessed due to the inadequacy of the model design used for dosing, that is, the 2×2 crossover model, discussed in Section 3.13.3.

Based on the results obtained for average bioequivalence and population bioequivalence, it can be indicated that bioequivalence studies can go a long way if applied by regulatory authorities in various countries on certifying drug formulations as bioequivalent or not, hence reducing healthcare costs. Generic drugs should be as safe and effective as brand name drugs, of which, to guarantee safety and effectiveness of generic drugs bioequivalence studies play a crucial role. It should be stressed that the quality of generic drugs should not be compromised, but should be similar to that of brand name drugs.

5.3 Conclusion

Many physicians and pharmacists have done research on average bioequivalence and the majority of them concur that there are limitations if drugs are certified as bioequivalent by assessing average bioequivalence only. They agree that it is not enough to show that drugs are average bioequivalent instead there is a need to investigate population and individual bioequivalence. As discussed above, population and individual bioequivalence have components which are not possessed by average bioequivalence which are important if drugs are to be certified as bioequivalent. ABE looks at only the comparison of means between the generic and brand name drug, while, PBE compare the means and variances of the test and reference formulation and has a property of prescribability. IBE compares means, within subject variances and subject by formulation interaction, hence has a property of switchability, where a patient can be safely and effectively transferred from one drug formulation to the other. PBE and IBE offer more components which are important when certifying drugs as being bioequivalent. But, there are challenges faced, especially when implementing IBE. Individual bioequivalence requires higher order designs and is expensive to carry out.

Bioequivalence studies are worthwhile in certifying drugs as bioequivalent, leading to people buying generic drugs at a lower cost as compared to brand name drugs, hence reducing healthcare costs. However, there are issues which make

people not to be completely satisfied when purchasing generic drugs. For example, bioequivalence procedures were not properly followed in the US in the late 1980s on certifying drugs as bioequivalent which led the public to develop lack of confidence on generic drugs. Other issues include:

- The development of generic drugs does not require large and extensive trials to be conducted on study subjects hence, generic drugs are still viewed as inferior compared to the brand name drugs;
- The use of the same acceptance limit(s) for all drugs by most regulatory authorities is questionable since some treatment formulations have a narrow or wide therapeutic range;
- Use of normal and healthy subjects, generally between 18 and 55 years is of concern since this sample age group cannot be representative of the whole population in various countries. Infants and elderly, say, above 70 years are likely to react differently when they take the same drugs with people who are in the interval 18 to 55 years;
- Packaging of the generic drugs (different from the brand name) can also be an issue to be concerned about, especially to the elderly (Meredith, 2003).

5.4 Recommendations

The use of generic drugs has generally gained momentum worldwide. As a result, bioequivalence studies still need to be improved so as to overcome some of the issues noted above, which lead some people to view generic drugs as inferior compared to brand name drugs. Issues which need to be addressed so as to improve bioequivalence include:

- Higher order crossover designs should be used because there are the ones needed to evaluate IBE. Drugs with high intra subject (within) or inter subject (between) variability which are not covered by ABE but are dealt with under IBE;
- Additional simulation assessment must be considered when evaluating the value of data collection period for PBE and IBE;

- The sensitivity of subject by formulation interaction to sample size and inherent variability of the compounds should be further explored through simulation studies.

The use of generic drugs all over the world will continue because generic drugs are cheaper, hence, reduces the cost of healthcare. Though generic drugs are cheaper, caution must always be taken when using them especially with regard to certain drug classes and patient populations.

References

Ahmad, T. and Yahya, S. (2015) *Bioequivalence in a 2×2 Crossover Design* pp. 1-15. www.createspace.com/4215274. [Accessed 2nd October 2015].

Becker, L.A. (2000) *Effect Size* pp. 1-14. http://web.uccs.edu/lbecker/Psy590/es.htm. [Accessed 29th September 2015].

Borgherini, G. (2003) The Bioequivalence and Therapeutic Efficacy of Generic Versus Brand name Psychoactive Drugs, *Clinical Therapeutics*, Vol. 25, Issue 6, pp. 1578-1592.

CDSCO (2005) *Guidelines for Bioavailability and Bioequivalence studies*, Central Drugs Standard Control Organization, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India, New Delhi, pp. 1-34.

Chen, M.L. (2007) Individual Bioequivalence-A regulatory Update, *Journal of Biopharmaceutical Statistics*, Vol. 7:1, pp. 5-11.

Chen, M.L., Patnaik, R., Hauck, W.W., Schuirmann, D.J., Hyslop, T. and Williams, R. (2000) An Individual Bioequivalence Criterion: Regulatory Considerations, *Statistics In Medicine*, Vol. 19; pp. 2821-2842.

Chen, M.L. and Lesko, L.J. (2001) Individual Bioequivalence Revisited, *Clin Pharmacokinet*, Vol. 40, No. 10, pp. 701-706.

Chernick, M.R. (2008) *Bootstrap Methods: A Guide for Practitioners and Researchers*, John Wiley and Sons, Inc; Hoboken; New Jersey, 369 pages.

Chinchilli, V.M. and Esinhart, J.D. (1996) Design and Analysis of Intra-Subject Variability in Crossover Experiments, *Statistics in Medicine*, Vol. 15, pp. 1619-1634.

Chow, S.C. and Liu, J.P. (2000) *Design and Analysis of Bioavailability and Bioequivalence*, Second Edition, New York, Chapman and Hall/CRC Biostatistics Series, 733 pages.

Chow, S.C. and Tse, S.K. (1990) Outlier Detection in Bioavailability and Bioequivalence Studies, *Statistics in Medicine*, Vol. 9, pp. 549-558.

Coe, R. It's the Effect Size, Stupid: What effect size is and why it is important: *Paper presented at the British Educational Research Association Annual Conference*, Exeter, September. 2002.

Cohen, J. (1992) A power primer, *Psychological Bulletin*, Vol. 112, No. 1, pp. 155-159.

Concordet, D. (2004) Statistics in Bioequivalence; *National Veterinary School*, pp. 1-89.

Crawshaw, J. and Chambers, J. (2001) *A concise course in Advanced Level Statistics*, Stanley Thornes Publishers Ltd, 686 pages.

Dighe, S.V. (1999) A Review of the Safety of Generic Drugs, *Elsevier Science Inc*, Vol. 31, Suppl 3A, pp. 235-245.

Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*, Chapman and Hall, London, 426 pages.

Elsayir, H.A. (2012) Computation and Interpretation of Effect Size in significance Test, *International Referred Journal of Engineering Science*, Vol. 1, No. 4, pp. 27-32.

Endrenyi, L., Amidon, G.L., Midha, K.K. and Skelly, J.P. (1998) Individual Bioequivalence: Attractive in Principle, Difficulty in Practice, *Pharmaceutical Research*, Vol. 15, No. 9, pp. 1321-1325.

FDA (1992) *Guidance on Statistical Procedures for Bioequivalence Using a Standard Two-Treatment Crossover Design* US Department of Health and Human

Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Rockville, Maryland, pp. 1-12.

FDA (1999) *Guidance for Industry: Population Pharmacokinetics*, US Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologies Evaluation and Research (CBER), Rockville, Maryland, pp. 1-31.

FDA (2001) *Guidance for Industry: Statistical Approaches to Bioequivalence*, US Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Rockville, Maryland pp. 1-48.

FDA (2003) *Guidance for Industry: Bioavailability and Bioequivalence Studies for Oral Administered Drug Products-General Considerations*, US Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Rockville, Maryland, pp. 1-22.

Galgate, U.C., Jamdade, V.R., Aute, P.P. and Chaudhari, P.D. (2013) Study on Requirements of Bioequivalence for Registration of Pharmaceutical Products in India, South Africa and Australia, *American Journal of PharmTech Research*, Vol. 3, No.1, pp. 290-305.

Galpin, J. (2013) *Statistical Research Design and Analysis* (STAT 4013 / 7009), School of Statistics and Actuarial Science, University of the Witwatersrand, 576 pages.

Greenblatt, D.J., Von Moltke, L.L., Harmatz, J.S. and Shader, R.I. (1998) *Pharmacokinetics, Pharmacodynamics, and Drug Disposition*. In:Davis KL, Charney D, Coyle JT, Nemeroff C, editors. Neuropsychopharmacology: the Fifth Generation of Progress Philadelphia: Lippincott, Williams and Wilkins, pp. 507-524.

Hauck, W.W. and Anderson, S. A. (1984) A new statistical procedure for testing equivalence in two-group bioavailablity trials, *J. Biopharm*, Vol. 12, pp. 83-91.

Japanese Guidelines (2001) *Clinical Pharmacokinetic Studies of Pharmaceuticals*, Ministry of Health, Labour and Welfare, pp. 1-23.

Jones, B. and Kenward, M.G. (2003) *Design and Analysis of Cross-Over Trials*, Chapman and Hall/CRC, 394 pages.

Kalpesh, S., Sokindra, K. and Kishore, M.K. (2013) Bridging the gap of Indian regulations and major global regulations for bioequivalence studies with emphasis on adaptive sequential design and two-stage bioequivalence studies, *Journal of Pharmaceutical and Scientific Innovation*, Vol. 2, No. 2, pp. 14-19.

Kamerow, D. (2011) The Pros and Cons of Generic Drugs, *BMI* Vol. 343, pp. 184.

Karasoy, D. and Daghan, G. (2012) Examination of Outliers in Bioequivalence Studies, *Clinical Psychopharmacology*, Vol. 22, no. 4, pp. 307-312.

Kelly, K. and Preacher, K.J. (2012) On Effect Size, *Psychological Methods*, Vol. 17, No. 2, pp. 137-152.

Lakens, D. (2013) Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs, *Frontiers in PSYCHOLOGY*, Vol. 4, Article 863, pp. 1-12.

Liu, J.P. and Weng, C.S. (1991) Detection of outlying data in bioavalability and bioequivalence studies, *Statistics in Medicine*, Vol. 10, pp. 1375-1389.

Lopes, N. (2009) Average, Population and Individual Bioequivalence, *SA Pharmaceutical Journal*, Vol. 10, pp. 46-48.

Mastan, S., Latha, T.B. and Ajay, S. (2011) The Basic Regulatory Considerations and Prospects for Conducting Bioavailability/Bioequivalence (BA/BE) Studies-An Overview, *Comparative Effectiveness Research*, Vol. 1, pp. 1-25.

McCarthy, W.F. and Guo, N. (2008) *Statistical Approach to Establishing Bioequivalence*, Maryland and Medical Research Institute, Baltimore Maryland, pp. 1-12.

Medicines Control Council of South Africa (2003) *Guideline on Application and Information Requirements-Veterinary Medicines*, pp. 1-47.

Meredith, P. (2003) The Bioequivalence and Other Unresolved Issues in Generic Drug Substitution, *Clinical Therapeutics*, Vol. 25, No. 11, pp. 2875-2890.

Moony, C.Z. and Duval, R.D. Bootstrapping. A nonparametric approach to statistical inference, *General Statistics Seminar*, University of Gronigen, October. 2012.

Microsoft® Office Excel® (2010). Download.com.CBS.

Midha, K.K. and Mckay, G. (2009) Bioequivalence; Its History, Practice, and Future, *The American Association of Pharmaceutical Scientists*, Vol. 11, No. 4, pp. 664-670.

Nandy, K. (2012) *Understanding and Quantifying Effect Sizes*, Translational Science Section, School of Nursing, Department of Biostatistics, School of Public Health, University of California, Los Angeles, pp. 1-41.

Patterson, S. (2010) *Bioequivalence and Statistics in clinical Pharmacology*, Villanova University, Summer III, pp. 1-110. http://www.homepp.villanova.edu/scott.patterson. [Accessed 4th December 2013].

Peer, A.V. (2007) Basic Concepts of Pharmacokinetics, *Clinical Pharmacology*, Gent, pp. 1-71.

Plets, M. and Strominger, J. Power Trip: A Road Map of PROC POWER, *Paper S1*, June 2013, pp. 1-4.

Polli, J.E. (2008) In Vitro Studies are sometimes Better than Conventional Human Pharmacokinetic In Vivo Studies in Assessing Bioequivalence of Immediate-

Release Solid Oral Dosage Forms, *The American Association of Pharmaceutical Scientists*, Vol. 10, Issue 2, pp. 289-299.

Qayyum, A. (2012) *Bioequivalence Studies, Readings in Advanced Pharmacokinetics-Theory, Methods and Applications*, Dr Ayman Noreddin (Ed), pp. 1-15. http://www.intechopen.com/books/readings-in-advanced-pharmacokinetics-theory-methods-and-applications/bioequivalence-studies. [Accessed 12th August 2013].

Ramsoy, D. and Elkum, N. (2004) A Comparison of Four Different Methods for Outlier Detection in Bioequivalence Studies, *Journal of Biopharmaceutical Statistics*, Vol.15, No.1, pp. 43-52.

Rani, S. and Pargal, A. (2004) Bioequivalence: An Overview of Statistical Concepts, *Educational Forum, Indian Journal Pharmacol*, Vol. 36, Issue 4, pp. 209-216.

Rice, K. and Thornotn, T. (2013) *More Loops, Control Structures and Bootstrapping*, University of Washington, pp. 1-23.

Rochowicz, J.A.Jr. (2011) Bootstrapping Analysis, Inferential Statistics and EXCEL, *Spreadsheets in Education (eJSiE)*, Vol. 4, Issue 3, Article 4, pp. 1-25. http://epublications.bond.edu.au/ejsie. [Accessed 15th August 2013].

SADC (2007) *SADC Guideline for Bioavailability and Bioequivalence*, pp. 1-25.

SAS® *Enterprise Guide Version 7.1*. (2014) SAS Institute Inc. Cary. NC. USA.

Schall, R. and Endrenyi, L. (2010) Bioequivalence: Tried and Tested, *Cardiovascular Journal of Africa*, Vol. 21, No. 2, pp. 69-71.

Schall, R., Ring, A. and Endrenyi, L. (2010) Residuals and Outliers in Replicate Design Crossover Studies, *Journal of Biopharmaceutical Studies*, Vol. 20, No. 4, pp. 835 – 849.

Schmidheiny, K. (2012) The Bootstrap, *Short Guides to Microeconometrics*, Spring, pp. 1-10.

Schuirmann, D.J. (1987) A comparison of two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability, *J. Pharmacokin Biopharm*, Vol. 15, pp. 657-680.

Schutz, H. Non-compartmental Analysis (NCA) in PK, PK-Based Design, Bioequivalence and Bioavailability, *Pre-conference* workshop, Budapest, 11 May 2009.

Schutz, H. Statistical Analysis of BE Data, Bioequivalence Assessment of Oral Dosage Forms *Basic Concepts and Practical Application Conference*, Leuven, June. 2013.

Senn, S. (2001) Cross-Over Trials in Drug Development: Theory and Practice, *Journal of Statistical Planning and Inference*, Vol. 96, pp. 96-40.

Steyn, H.S. (1999) *Manual for the determination of Effect Size indices and Practical Significance*, Statistical Consultation Services, North West University, Potchefstroom Campus, 211 pages.

Sun, P. (2010) Using SAS Proc Power to Perform Model-based Power Analysis for Clinical Pharmacology Studies, *Pharma SUG 2010*, Vol 5, pp. 1-11.

Tamboli, A.M., Todkar, P. and Sayyad, F.J. (2010) An Overview on Bioequivalence: Regulatory Consideration for Generic Drug Products, *Journal of Bioequivalence and Bioavailability*, Vol 2, No 4, pp. 86-92.

The European Agency for the Evaluation of Medicinal Products. (2001) *Guidelines for the Conduct of Bioequivalence Studies for Veterinary Medicinal Products*, pp. 1-11. http://www.ec.europa.eu/health/files/eudralex/ [Accessed 24th October 2016].

Wang, D. and Bakhai, A. (2006) *Clinical Trials-A Practical Guide to Design, Analysis and Reporting*, London, Remedica, 480 pages.

Wang, W. and Chow, S.C. (2003) Examining outlying subjects and outlying records in bioequivalence trials, *Journal of Biopharmaceutical Statistics*, Vol. 13, pp. 43-56.

Warmbrod, J.R. Conducting, Interpreting and Reporting Quantitative research *Research Pre-Session , National Agricultural Education Research Conference*, New Orleans, Los Angeles December 2001.

Westlake, W.J (1976) Symmetrical Confidence Intervals for Bioequivalence Trials, *Biometrics*, Vol. 32, No. 4, pp. 741-744.

Wilson, D.B. Calculating Effect-Sizes *The Campbell Foundation*, George Mason University, August 2011.

Appendices

Appendix 1 Concentration Time Profiles for Subject (Sheep) and Period after dosing with Reference and Test Formulations.

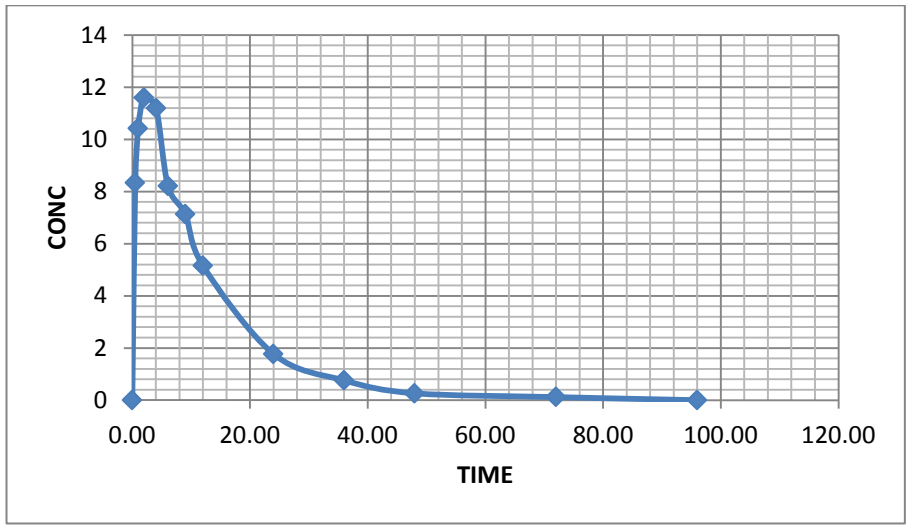


Figure 1: Fit Plot for Concentration Time Profile for Sheep 1, Sequence 2, Test Treatment in Period 2

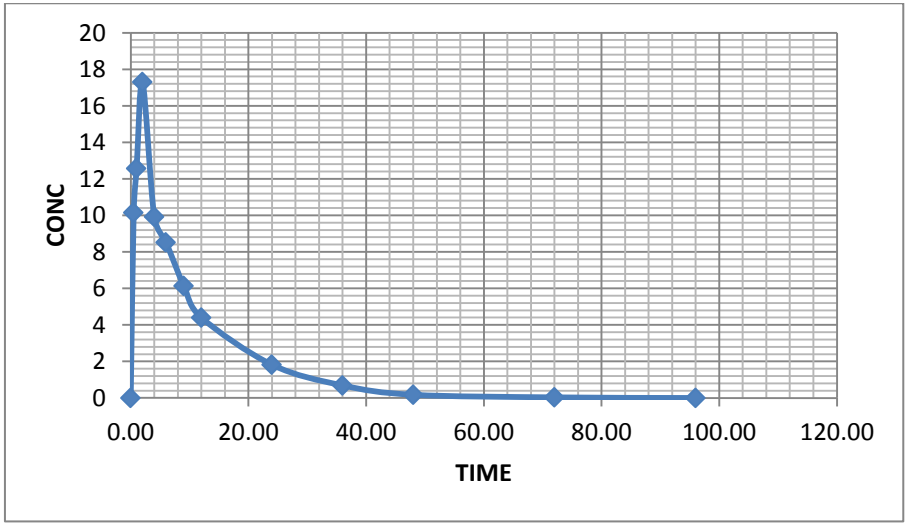


Figure 2: Fit Plot for Concentration Time Profile for Sheep 1, Sequence 2, Reference Treatment in Period 1

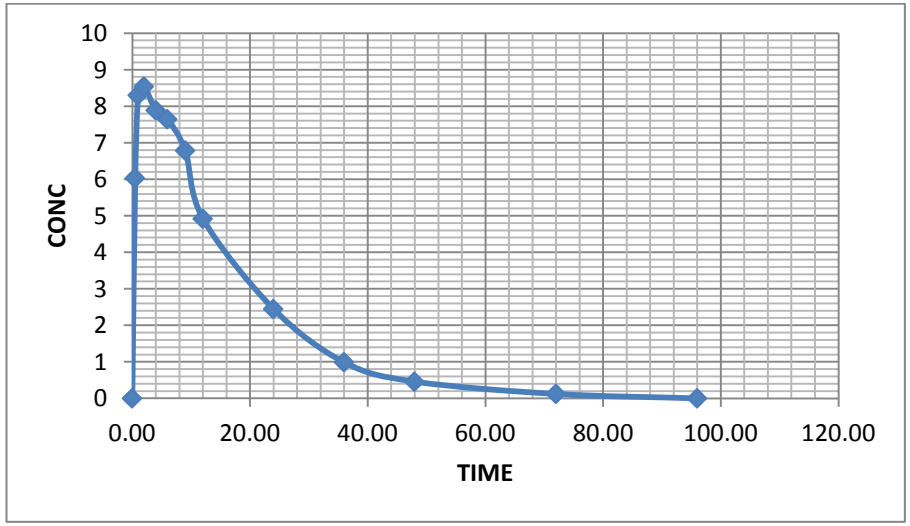


Figure 3: Fit Plot for Concentration Time Profile for Sheep 2, Sequence 2, Reference Treatment in Period 1

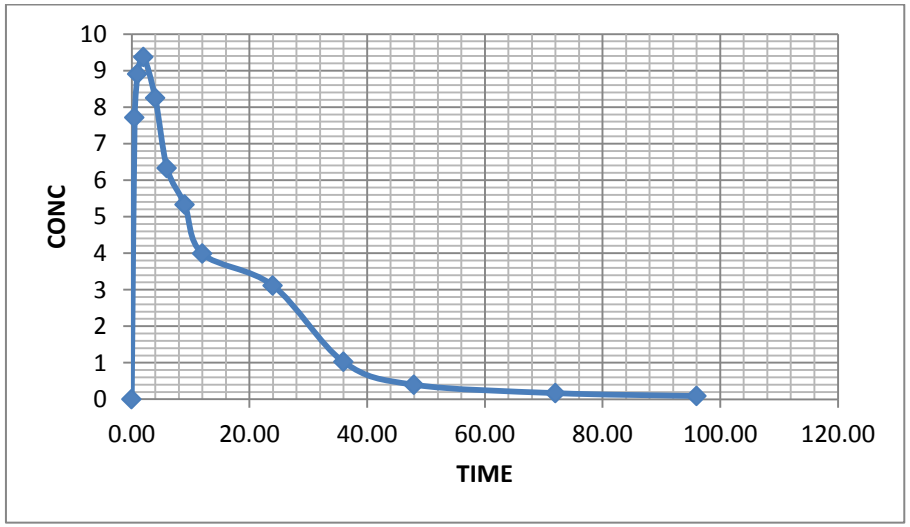


Figure 4: Fit Plot for Concentration Time Profile for Sheep 2, Sequence 2, Test Treatment in Period 2

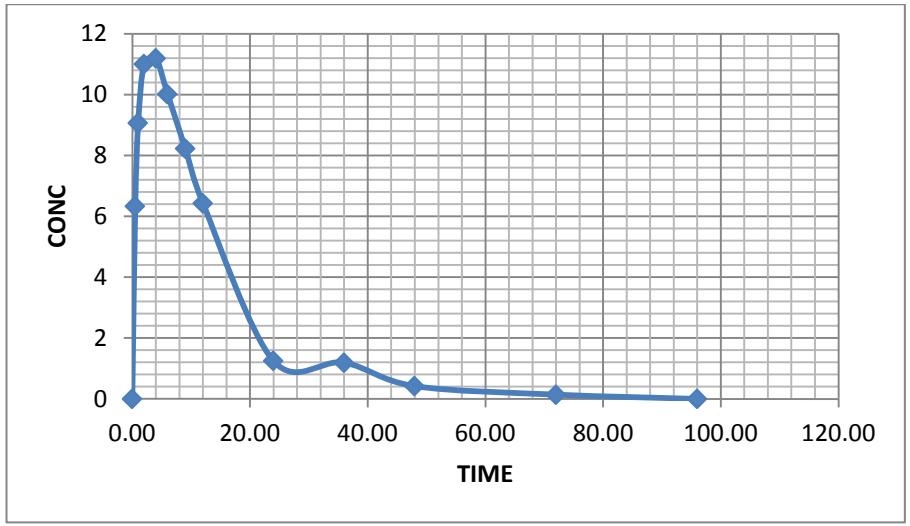


Figure 5: Fit Plot for Concentration Time Profile for Sheep 3, Sequence 2, Test Treatment in Period 2

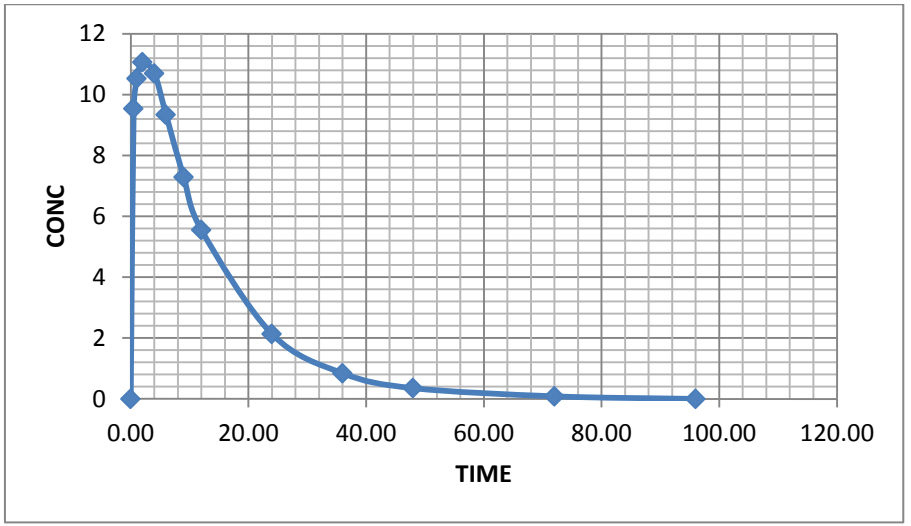


Figure 6: Fit Plot for Concentration Time Profile for Sheep 3, Sequence 2, Reference Treatment in Period 1

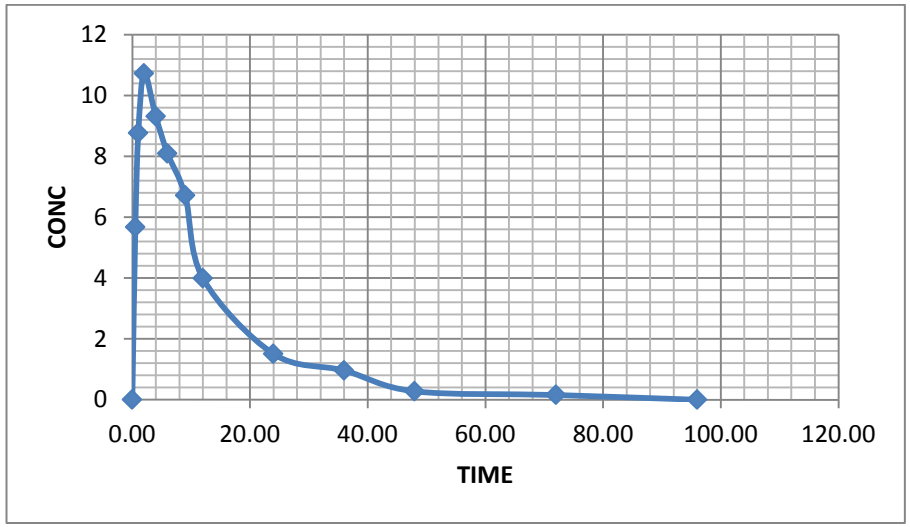


Figure 7: Fit Plot for Concentration Time Profile for Sheep 4, Sequence 2, Reference Treatment in Period 1

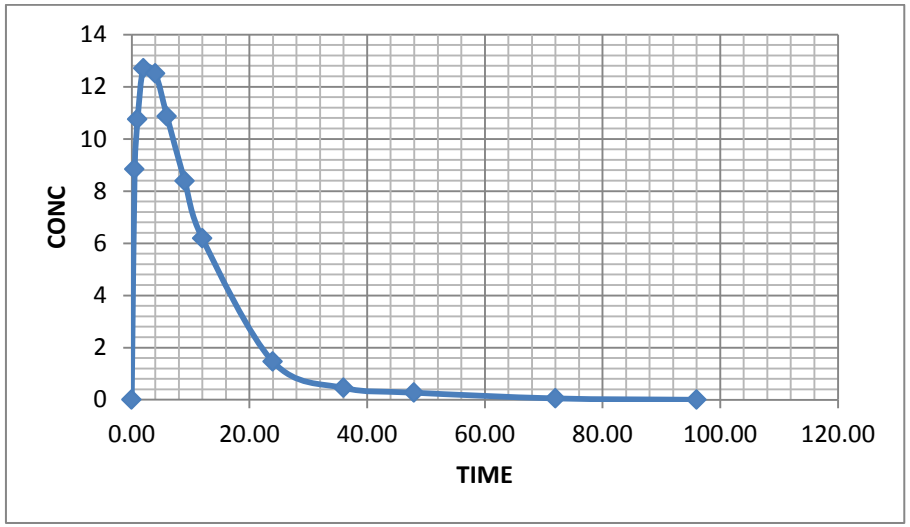


Figure 8: Fit Plot for Concentration Time Profile for Sheep 4, Sequence 2, Test Treatment in Period 2

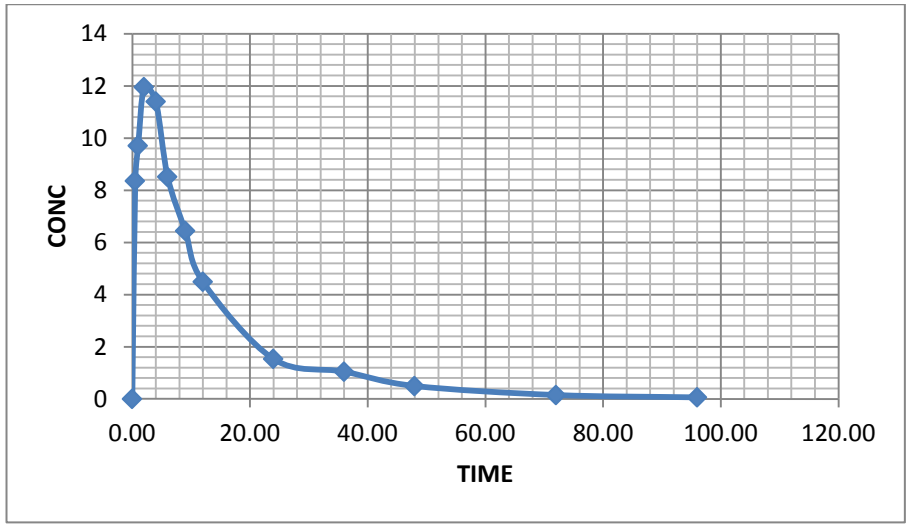


Figure 9: Fit Plot for Concentration Time Profile for Sheep 5, Sequence 1, Reference Treatment in Period 2

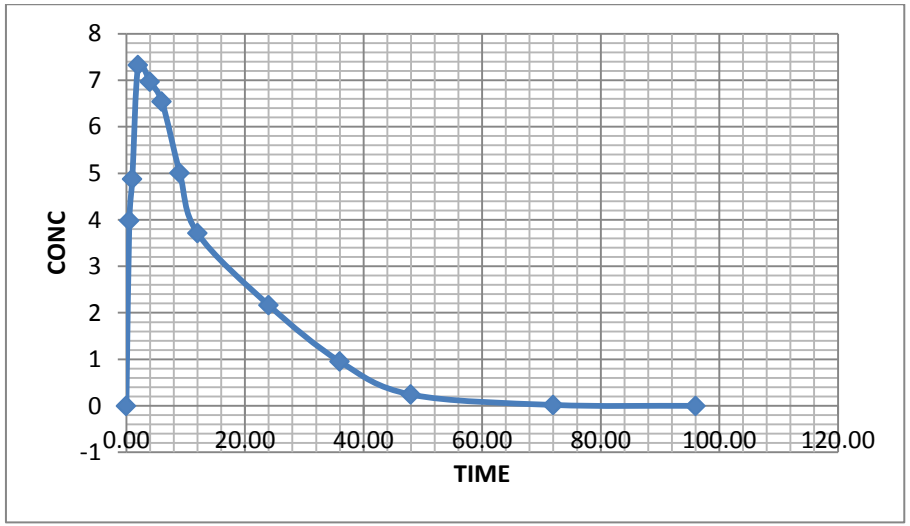


Figure 10: Fit Plot for Concentration Time Profile for Sheep 5, Sequence 1, Test Treatment in Period 1

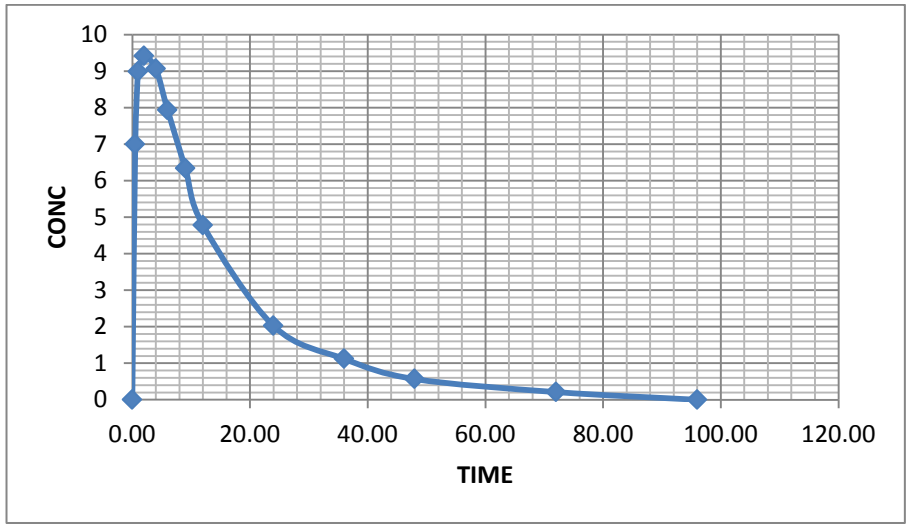


Figure 11: Fit Plot for Concentration Time Profile for Sheep 6, Sequence 1, Test Treatment in Period 1

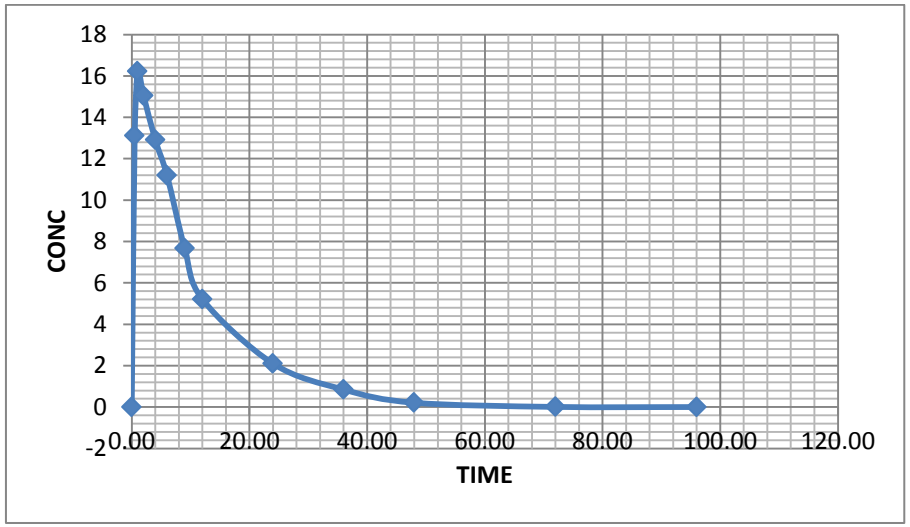


Figure 12: Fit Plot for Concentration Time Profile for Sheep 6, Sequence 1, Reference Treatment in Period 2

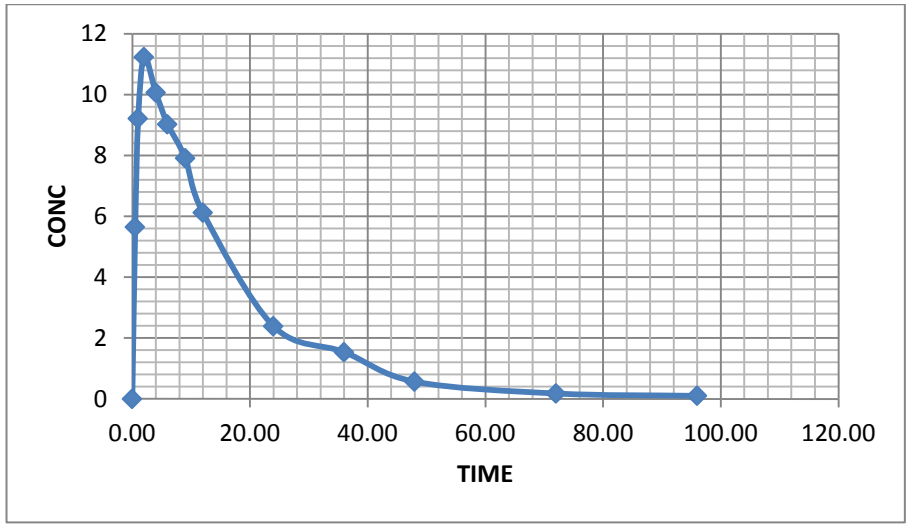


Figure 13: Fit Plot for Concentration Time Profile for Sheep 7, Sequence 2, Test Treatment in Period 1

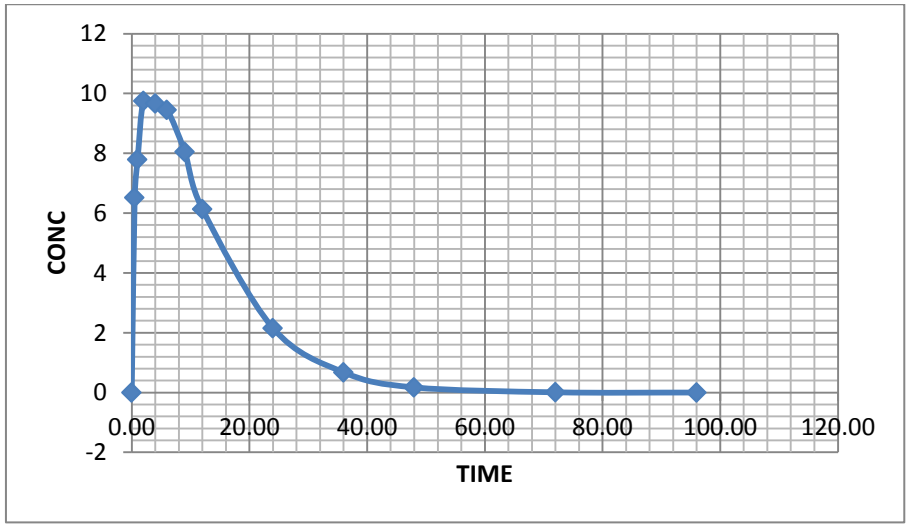


Figure 14: Fit Plot for Concentration Time Profile for Sheep 7, Sequence 2, Test Treatment in Period 1

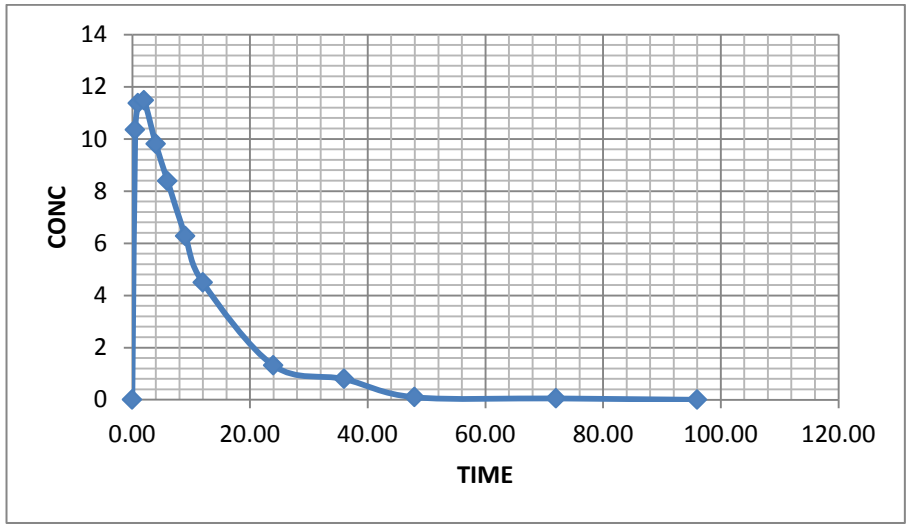


Figure 15: Fit Plot for Concentration Time Profile for Sheep 8, Sequence 2, Reference Treatment in Period 1

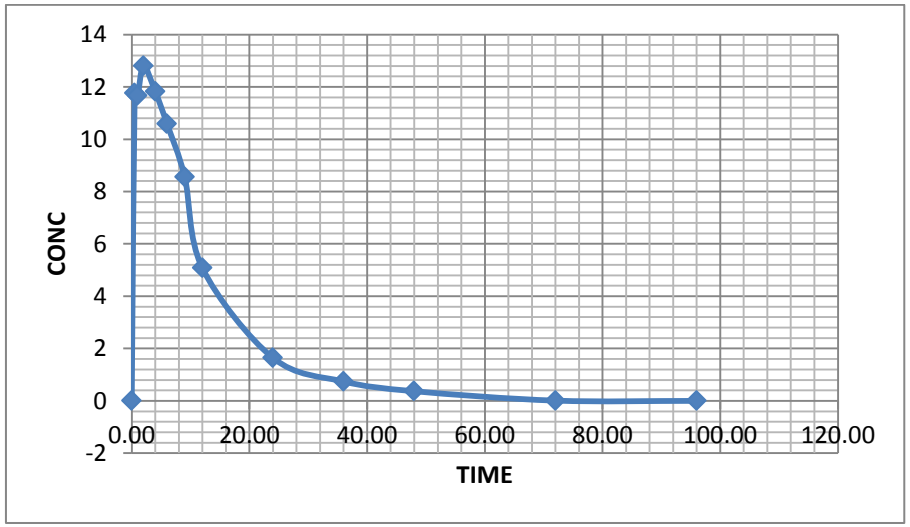


Figure 16: Fit Plot for Concentration Time Profile for Sheep 8, Sequence 2, Test Treatment in Period 2

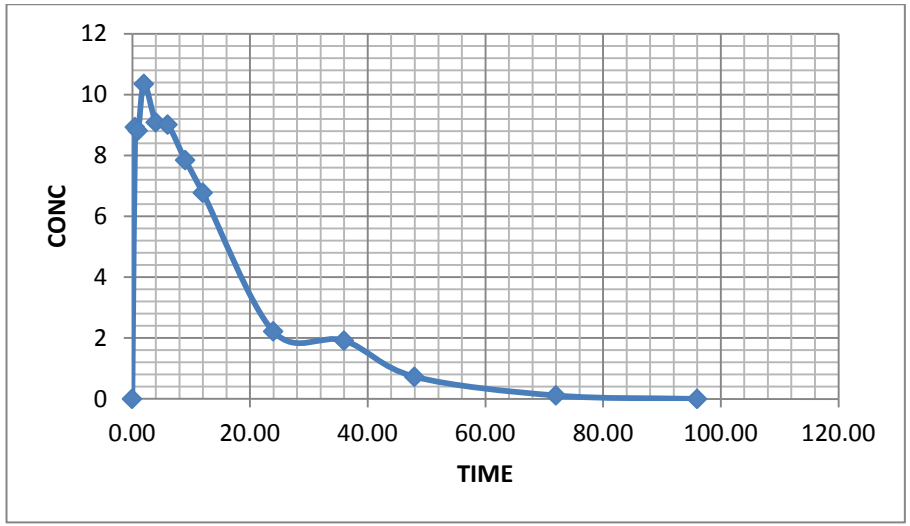


Figure 17: Fit Plot for Concentration Time Profile for Sheep 9, Sequence 1, Reference Treatment in Period 2

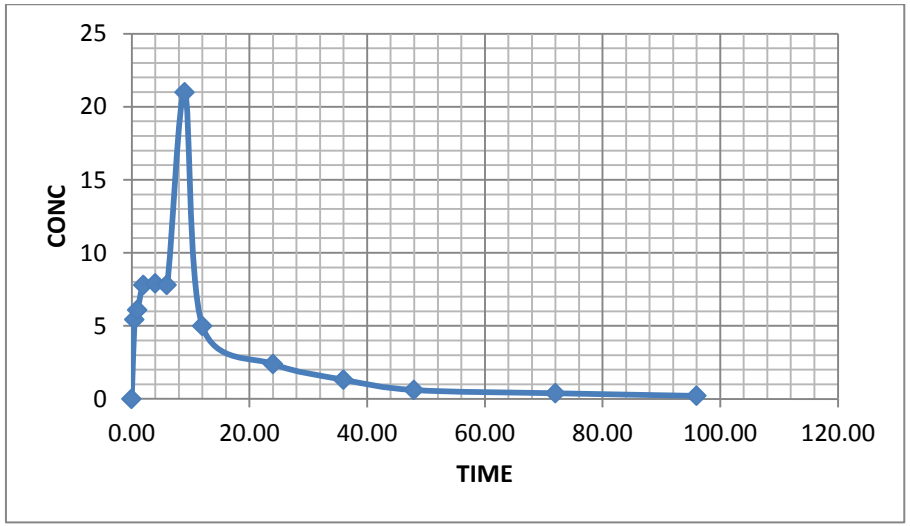


Figure 18: Fit Plot for Concentration Time Profile for Sheep 9, Sequence 1, Test Treatment in Period 1

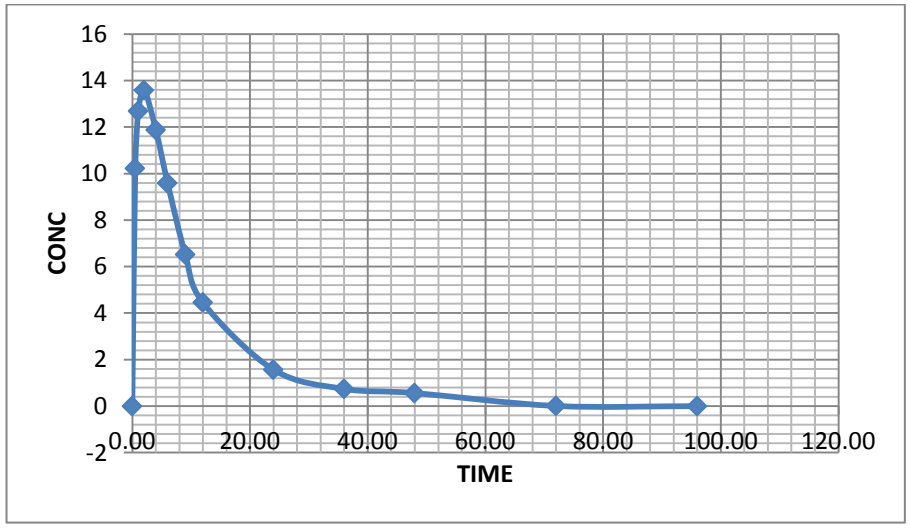


Figure 19: Fit Plot for Concentration Time Profile for Sheep 10, Sequence 1, Test Treatment in Period 1

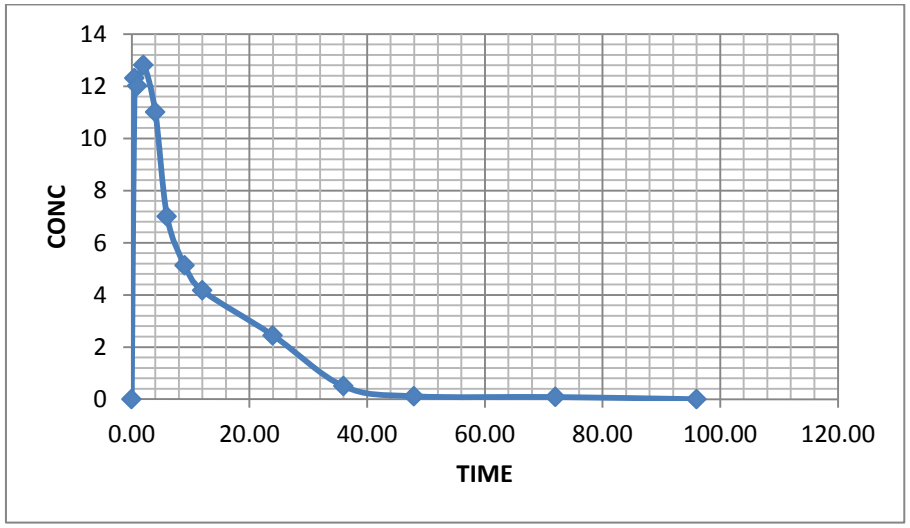


Figure 20: Fit Plot for Concentration Time Profile for Sheep 10, Sequence 1, Reference Treatment in Period 2

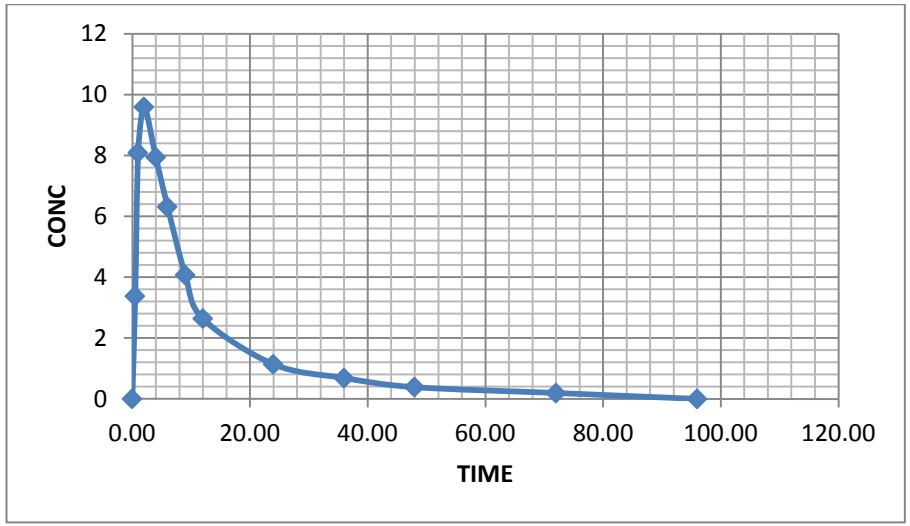


Figure 21: Fit Plot for Concentration Time Profile for Sheep 11, Sequence 1, Test Treatment in Period 1

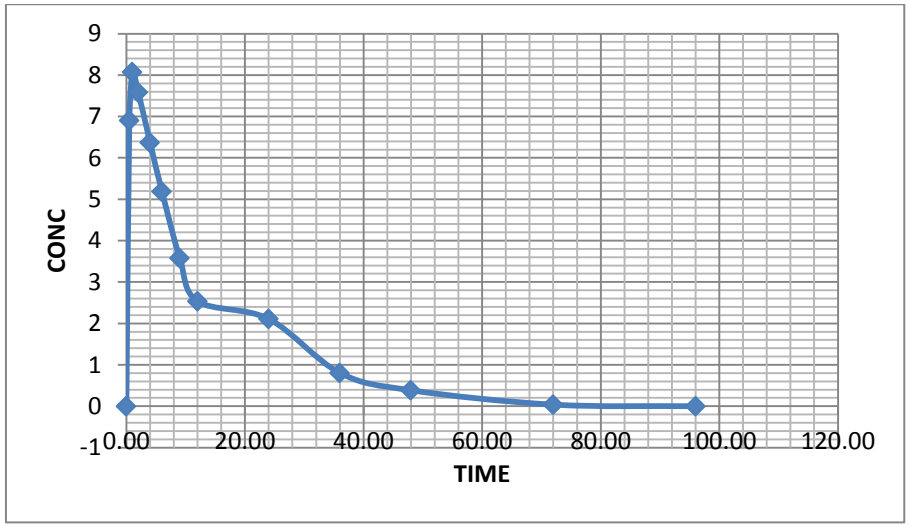


Figure 22: Fit Plot for Concentration Time Profile for Sheep 11, Sequence 1, Reference Treatment in Period 2

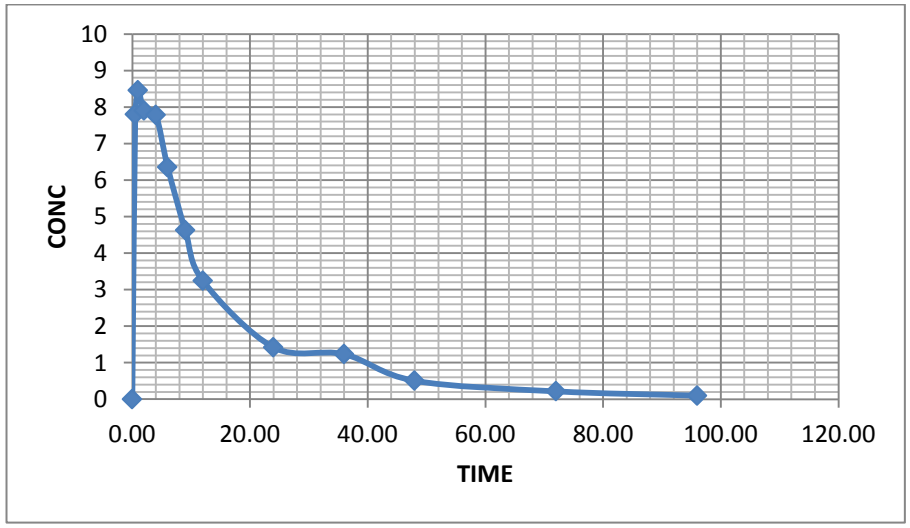


Figure 23: Fit Plot for Concentration Time Profile for Sheep 12, Sequence 1, Test Treatment in Period 1

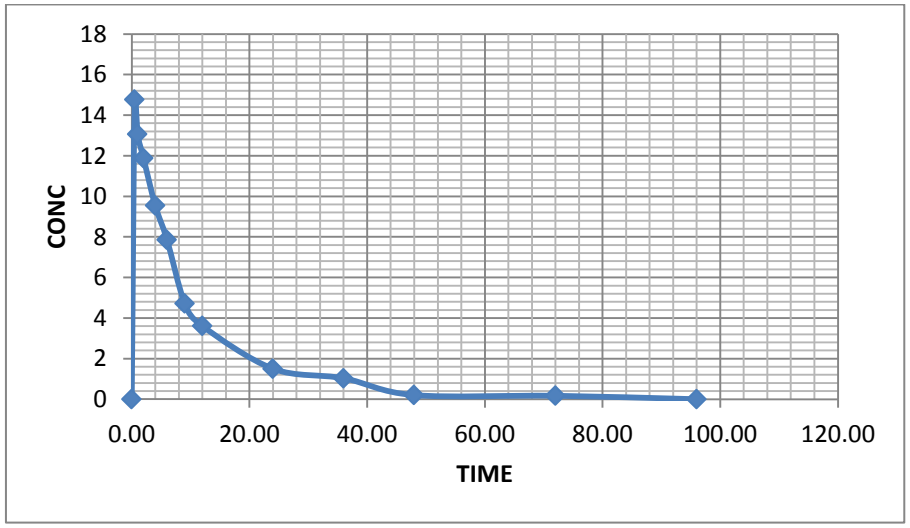


Figure 24: Fit Plot for Concentration Time Profile for Sheep 12, Sequence 1, Reference Treatment in Period 2

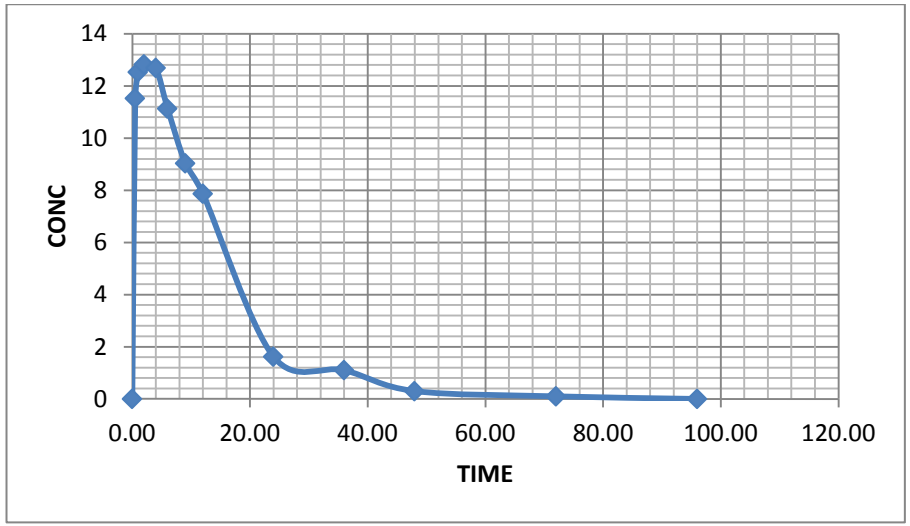


Figure 25: Fit Plot for Concentration Time Profile for Sheep 13, Sequence 2, Test Treatment in Period 2

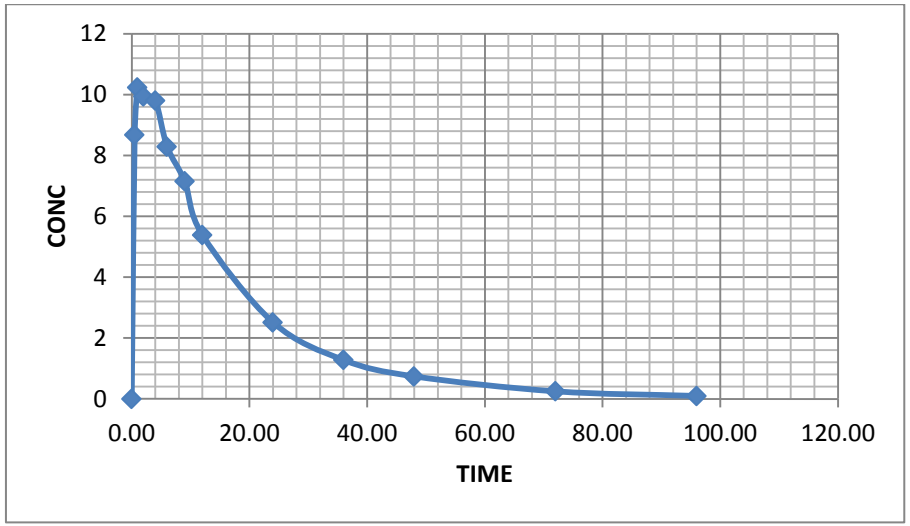


Figure 26: Fit Plot for Concentration Time Profile for Sheep 13, Sequence 2, Reference Treatment in Period 1

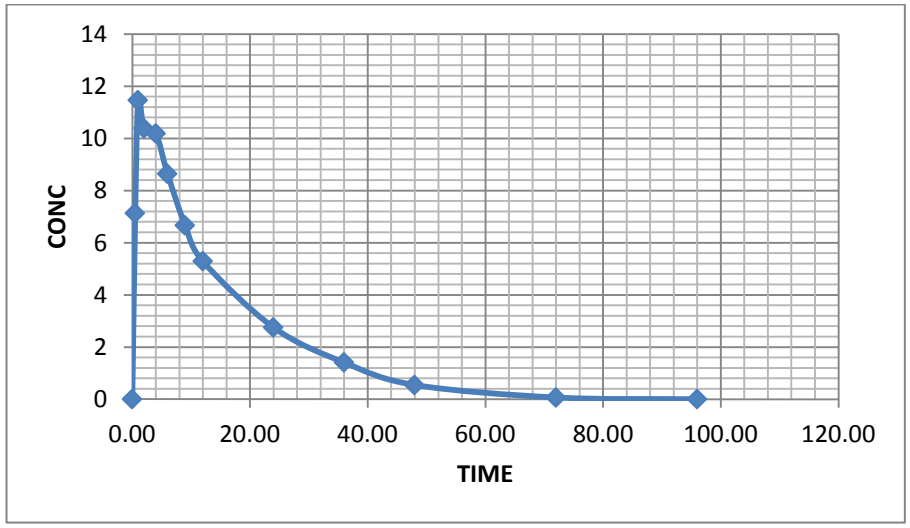


Figure 27: Fit Plot for Concentration Time Profile for Sheep 14, Sequence 1, Test Treatment in Period 1

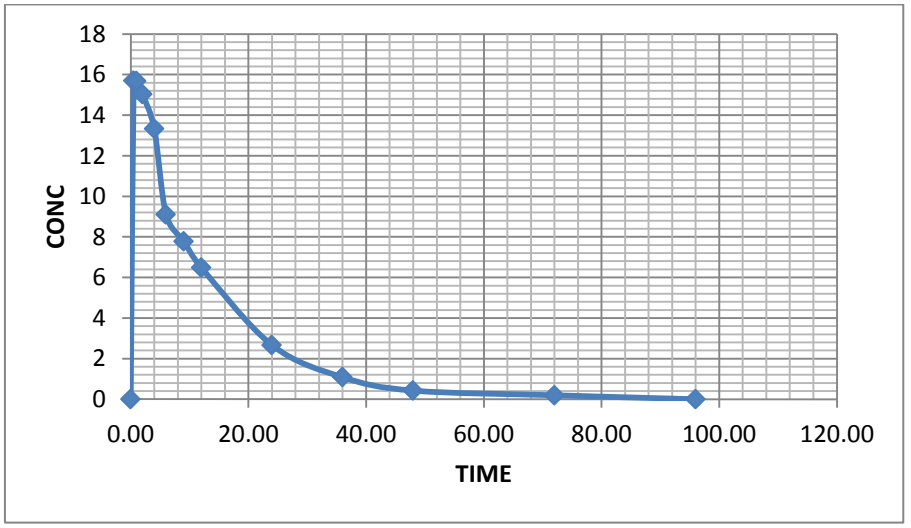


Figure 28: Fit Plot for Concentration Time Profile for Sheep 14, Sequence 1, Reference Treatment in Period 2

Appendix 2 Calculating *AUC* from time zero to time *t*.

Actual time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}		Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	1	0	T	NA	NA	NA		0.00	1	0	R	NA	NA	NA
0.50	1	8.3332	T	4.1666	0.50	2.0833		0.50	1	10.1554	R	5.0777	0.50	2.53885
1.00	1	10.4159	T	9.37455	0.50	4.687275		1.00	1	12.5728	R	11.3641	0.50	5.68205
2.00	1	11.5974	T	11.00665	1.00	11.00665		2.00	1	17.303	R	14.9379	1.00	14.9379
4.05	1	11.1884	T	11.3929	2.05	23.35545		4.00	1	9.9144	R	13.6087	2.00	27.2174
6.07	1	8.204	T	9.6962	2.02	19.554		6.00	1	8.5193	R	9.21685	2.00	18.4337
9.00	1	7.1336	T	7.6688	2.93	22.49515		9.03	1	6.1342	R	7.32675	3.03	22.22448
12.00	1	5.1534	T	6.1435	3.00	18.4305		12.00	1	4.4043	R	5.26925	2.97	15.63211
24.00	1	1.7674	T	3.4604	12.00	41.5248		24.00	1	1.8197	R	3.112	12.00	37.344
36.00	1	0.7566	T	1.262	12.00	15.144		36.00	1	0.687	R	1.25335	12.00	15.0402
48.00	1	0.2632	T	0.5099	12.00	6.1188		48.03	1	0.1749	R	0.43095	12.03	5.185765
72.00	1	0.1173	T	0.19025	24.00	4.566		72.00	1	0.0336	R	0.10425	23.97	2.498525
96.00	1	0	T	0.05865	24.00	1.4076		96.02	1	0	R	0.0168	24.02	0.40348
						170.3735								167.1385

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	2	0	R	NA	NA	NA
0.50	2	6.0324	R	3.0162	0.50	1.5081
1.00	2	8.3025	R	7.16745	0.50	3.583725
2.00	2	8.5327	R	8.4176	1.00	8.4176
4.00	2	7.8854	R	8.20905	2.00	16.4181
6.00	2	7.6404	R	7.7629	2.00	15.5258
9.00	2	6.7815	R	7.21095	3.00	21.63285
12.00	2	4.9191	R	5.8503	3.00	17.5509
24.00	2	2.4476	R	3.68335	12.00	44.2002
36.00	2	0.9889	R	1.71825	12.00	20.619
48.00	2	0.458	R	0.72345	12.00	8.6814
72.00	2	0.1206	R	0.2893	24.00	6.9432
96.00	2	0	R	0.0603	24.00	1.4472
						166.5281

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	2	0	T	NA	NA	NA
0.50	2	7.7087	T	3.85435	0.50	1.927175
1.00	2	8.9049	T	8.3068	0.50	4.1534
2.00	2	9.3668	T	9.13585	1.00	9.13585
4.00	2	8.2435	T	8.80515	2.00	17.6103
6.00	2	6.3296	T	7.28655	2.00	14.5731
9.00	2	5.3302	T	5.8299	3.00	17.4897
12.00	2	3.9859	T	4.65805	3.00	13.97415
24.00	2	3.113	T	3.54945	12.00	42.5934
36.00	2	1.0224	T	2.0677	12.00	24.8124
48.00	2	0.3971	T	0.70975	12.00	8.517
72.00	2	0.1688	T	0.28295	24.00	6.7908
96.00	2	0.0885	T	0.12865	24.00	3.0876
						164.6649

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	3	0	T	NA	NA	NA
0.50	3	6.3276	T	3.1638	0.50	1.5819
1.00	3	9.0636	T	7.6956	0.50	3.8478
2.00	3	11.0009	T	10.03225	1.00	10.03225
4.00	3	11.1924	T	11.09665	2.00	22.1933
6.00	3	10.0088	T	10.6006	2.00	21.2012
9.00	3	8.2297	T	9.11925	3.00	27.35775
12.00	3	6.4307	T	7.3302	3.00	21.9906
24.00	3	1.248	T	3.83935	12.00	46.0722
36.00	3	1.1917	T	1.21985	12.00	14.6382
48.00	3	0.4199	T	0.8058	12.00	9.6696
72.00	3	0.1407	T	0.2803	24.00	6.7272
96.00	3	0	T	0.07035	24.00	1.6884
						187.0004

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	3	0	R	NA	NA	NA
0.50	3	9.5422	R	4.7711	0.50	2.38555
1.03	3	10.5317	R	10.03695	0.53	5.35304
2.00	3	11.0619	R	10.7968	0.97	10.43691
4.00	3	10.7004	R	10.88115	2.00	21.7623
6.00	3	9.3407	R	10.02055	2.00	20.0411
9.00	3	7.2836	R	8.31215	3.00	24.93645
12.00	3	5.5551	R	6.41935	3.00	19.25805
24.00	3	2.1309	R	3.843	12.00	46.116
36.00	3	0.8436	R	1.48725	12.00	17.847
48.00	3	0.3536	R	0.5986	12.00	7.1832
72.00	3	0.0852	R	0.2194	24.00	5.2656
96.00	3	0	R	0.0426	24.00	1.0224
						181.6076

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	4	0	R	NA	NA	NA
0.52	4	5.6722	R	2.8361	0.52	1.465318
1.02	4	8.761	R	7.2166	0.50	3.6083
2.00	4	10.7164	R	9.7387	0.98	9.576388
4.00	4	9.3046	R	10.0105	2.00	20.021
6.00	4	8.0849	R	8.69475	2.00	17.3895
9.00	4	6.7065	R	7.3957	3.00	22.1871
12.00	4	3.9893	R	5.3479	3.00	16.0437
24.00	4	1.5077	R	2.7485	12.00	32.982
36.00	4	0.9594	R	1.23355	12.00	14.8026
48.00	4	0.2773	R	0.61835	12.00	7.4202
72.00	4	0.1561	R	0.2167	24.00	5.2008
96.00	4	0	R	0.07805	24.00	1.8732
						152.5701

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	4	0	T	NA	NA	NA
0.50	4	8.84	T	4.42	0.50	2.21
1.00	4	10.7492	T	9.7946	0.50	4.8973
2.00	4	12.6993	T	11.72425	1.00	11.72425
4.00	4	12.5011	T	12.6002	2.00	25.2004
6.00	4	10.8577	T	11.6794	2.00	23.3588
9.00	4	8.3821	T	9.6199	3.00	28.8597
12.00	4	6.1864	T	7.28425	3.00	21.85275
24.00	4	1.4588	T	3.8226	12.00	45.8712
36.00	4	0.4553	T	0.95705	12.00	11.4846
48.00	4	0.2713	T	0.3633	12.00	4.3596
72.00	4	0.0517	T	0.1615	24.00	3.876
96.03	4	0	T	0.02585	24.03	0.621262
						184.3159

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	5	0	R	NA	NA	NA
0.50	5	8.3506	R	4.1753	0.50	2.08765
1.00	5	9.7089	R	9.02975	0.50	4.514875
2.00	5	11.9468	R	10.82785	1.00	10.82785
4.00	5	11.3999	R	11.67335	2.00	23.3467
6.00	5	8.5153	R	9.9576	2.00	19.9152
9.00	5	6.4408	R	7.47805	3.00	22.43415
12.00	5	4.486	R	5.4634	3.00	16.3902
24.00	5	1.5305	R	3.00825	12.00	36.099
36.00	5	1.0458	R	1.28815	12.00	15.4578
48.00	5	0.4962	R	0.771	12.00	9.252
72.00	5	0.1494	R	0.3228	24.00	7.7472
96.00	5	0.0597	R	0.10455	24.00	2.5092
						170.5818

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	5	0	T	NA	NA	NA
0.50	5	3.9791	T	1.98955	0.50	0.994775
1.00	5	4.8763	T	4.4277	0.50	2.21385
2.00	5	7.3284	T	6.10235	1.00	6.10235
4.00	5	6.9689	T	7.14865	2.00	14.2973
6.00	5	6.5418	T	6.75535	2.00	13.5107
9.00	5	5.0062	T	5.774	3.00	17.322
12.00	5	3.7108	T	4.3585	3.00	13.0755
24.00	5	2.1651	T	2.93795	12.00	35.2554
36.00	5	0.9521	T	1.5586	12.00	18.7032
48.00	5	0.2425	T	0.5973	12.00	7.1676
72.00	5	0.0195	T	0.131	24.00	3.144
96.00	5	0	T	0.00975	24.00	0.234
						132.0207

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	6	0	T	NA	NA	NA
0.50	6	6.9937	T	3.49685	0.50	1.748425
1.00	6	8.9846	T	7.98915	0.50	3.994575
2.00	6	9.4104	T	9.1975	1.00	9.1975
4.00	6	9.0616	T	9.236	2.00	18.472
6.00	6	7.9323	T	8.49695	2.00	16.9939
9.00	6	6.341	T	7.13665	3.00	21.40995
12.00	6	4.7806	T	5.5608	3.00	16.6824
24.00	6	2.0218	T	3.4012	12.00	40.8144
36.00	6	1.1188	T	1.5703	12.00	18.8436
48.00	6	0.5665	T	0.84265	12.00	10.1118
72.00	6	0.2103	T	0.3884	24.00	9.3216
96.00	6	0	T	0.10515	24.00	2.5236
						170.1138

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	6	0	R	NA	NA	NA
0.50	6	13.105	R	6.5525	0.50	3.27625
1.00	6	16.2239	R	14.66445	0.50	7.332225
2.02	6	15.0403	R	15.6321	1.02	15.89264
4.00	6	12.9108	R	13.97555	1.98	27.71817
6.00	6	11.1877	R	12.04925	2.00	24.0985
9.00	6	7.6651	R	9.4264	3.00	28.2792
12.00	6	5.213	R	6.43905	3.00	19.31715
24.00	6	2.1001	R	3.65655	12.00	43.8786
36.00	6	0.845	R	1.47255	12.00	17.6706
48.00	6	0.209	R	0.527	12.00	6.324
72.00	6	0.0066	R	0.1078	24.00	2.5872
96.00	6	0	R	0.0033	24.00	0.0792
						196.4537

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	7	0	T	NA	NA	NA
0.50	7	5.6495	T	2.82475	0.50	1.412375
1.00	7	9.2156	T	7.43255	0.50	3.716275
2.00	7	11.2346	T	10.2251	1.00	10.2251
4.00	7	10.0604	T	10.6475	2.00	21.295
6.00	7	9.0281	T	9.54425	2.00	19.0885
9.00	7	7.9102	T	8.46915	3.00	25.40745
12.00	7	6.1167	T	7.01345	3.00	21.04035
24.00	7	2.3893	T	4.253	12.00	51.036
36.00	7	1.5392	T	1.96425	12.00	23.571
48.00	7	0.5665	T	1.05285	12.00	12.6342
72.00	7	0.1782	T	0.37235	24.00	8.9364
96.00	7	0.1005	T	0.13935	24.00	3.3444
						201.7071

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	7	0	R	NA	NA	NA
0.50	7	6.5191	R	3.25955	0.50	1.629775
1.00	7	7.7943	R	7.1567	0.50	3.57835
2.00	7	9.7471	R	8.7707	1.00	8.7707
4.00	7	9.6688	R	9.70795	2.00	19.4159
6.00	7	9.4525	R	9.56065	2.00	19.1213
9.00	7	8.0507	R	8.7516	3.00	26.2548
12.00	7	6.1368	R	7.09375	3.00	21.28125
24.00	7	2.155	R	4.1459	12.00	49.7508
36.00	7	0.6776	R	1.4163	12.00	16.9956
48.00	7	0.1722	R	0.4249	12.00	5.0988
72.00	7	0.0079	R	0.09005	24.00	2.1612
96.00	7	0	R	0.00395	24.00	0.0948
						174.1533

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	8	0	R	NA	NA	NA
0.50	8	10.3496	R	5.1748	0.50	2.5874
1.00	8	11.3571	R	10.85335	0.50	5.426675
2.00	8	11.4669	R	11.412	1.00	11.412
4.00	8	9.798	R	10.63245	2.00	21.2649
6.00	8	8.3814	R	9.0897	2.00	18.1794
9.00	8	6.2741	R	7.32775	3.00	21.98325
12.00	8	4.4873	R	5.3807	3.00	16.1421
24.00	8	1.3102	R	2.89875	12.00	34.785
36.00	8	0.7934	R	1.0518	12.00	12.6216
48.00	8	0.0965	R	0.44495	12.00	5.3394
72.00	8	0.0507	R	0.0736	24.00	1.7664
96.00	8	0	R	0.02535	24.00	0.6084
						152.1165

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	8	0	T	NA	NA	NA
0.50	8	11.7661	T	5.88305	0.50	2.941525
1.00	8	11.6697	T	11.7179	0.50	5.85895
2.00	8	12.7964	T	12.23305	1.00	12.23305
4.00	8	11.8217	T	12.30905	2.00	24.6181
6.00	8	10.5832	T	11.20245	2.00	22.4049
9.00	8	8.5582	T	9.5707	3.00	28.7121
12.00	8	5.0771	T	6.81765	3.00	20.45295
24.00	8	1.6423	T	3.3597	12.00	40.3164
36.00	8	0.7412	T	1.19175	12.00	14.301
48.00	8	0.369	T	0.5551	12.00	6.6612
72.00	8	0.0099	T	0.18945	24.00	4.5468
96.00	8	0	T	0.00495	24.00	0.1188
						183.1658

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}		Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	9	0	R	NA	NA	NA		0.00	9	0	T	NA	NA	NA
0.50	9	8.9297	R	4.46485	0.50	2.232425		0.50	9	5.4319	T	2.71595	0.50	1.357975
1.00	9	8.8112	R	8.87045	0.50	4.435225		1.00	9	6.0926	T	5.76225	0.50	2.881125
2.00	9	10.3462	R	9.5787	1.00	9.5787		2.00	9	7.797	T	6.9448	1.00	6.9448
4.02	9	9.0897	R	9.71795	2.02	19.59787		4.03	9	7.9162	T	7.8566	2.03	15.97509
6.02	9	9.0147	R	9.0522	2.00	18.1044		6.00	9	7.8071	T	7.86165	1.97	15.46125
9.00	9	7.8452	R	8.42995	2.98	25.14935		9.00	9	21	T	14.40355	3.00	43.21065
12.00	9	6.7688	R	7.307	3.00	21.921		12.00	9	4.9914	T	12.9957	3.00	38.9871
24.00	9	2.2146	R	4.4917	12.00	53.9004		24.02	9	2.3927	T	3.69205	12.02	44.36613
36.00	9	1.9167	R	2.06565	12.00	24.7878		36.00	9	1.3136	T	1.85315	11.98	22.20691
48.00	9	0.7285	R	1.3226	12.00	15.8712		48.00	9	0.6073	T	0.96045	12.00	11.5254
72.00	9	0.1126	R	0.42055	24.00	10.0932		72.00	9	0.3871	T	0.4972	24.00	11.9328
96.00	9	0	R	0.0563	24.00	1.3512		96.00	9	0.209	T	0.29805	24.00	7.1532
						207.0228								222.0024

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	10	0	T	NA	NA	NA
0.50	10	10.2244	T	5.1122	0.50	2.5561
1.02	10	12.6899	T	11.45715	0.52	5.919528
2.02	10	13.5789	T	13.1344	1.00	13.1344
4.00	10	11.8779	T	12.7284	1.98	25.24466
6.00	10	9.5904	T	10.73415	2.00	21.4683
9.00	10	6.5224	T	8.0564	3.00	24.1692
12.00	10	4.4539	T	5.48815	3.00	16.46445
24.00	10	1.5653	T	3.0096	12.00	36.1152
36.00	10	0.7419	T	1.1536	12.00	13.8432
48.00	10	0.5626	T	0.65225	12.00	7.827
72.00	10	0.0075	T	0.28505	24.00	6.8412
96.00	10	0	T	0.00375	24.00	0.09
						173.6732

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	10	0	R	NA	NA	NA
0.50	10	12.3016	R	6.1508	0.50	3.0754
1.00	10	12.0091	R	12.15535	0.50	6.077675
2.00	10	12.8158	R	12.41245	1.00	12.41245
4.00	10	11.009	R	11.9124	2.00	23.8248
6.00	10	7.0098	R	9.0094	2.00	18.0188
9.00	10	5.1347	R	6.07225	3.00	18.21675
12.00	10	4.166	R	4.65035	3.00	13.95105
24.00	10	2.4456	R	3.3058	12.00	39.6696
36.00	10	0.4982	R	1.4719	12.00	17.6628
48.00	10	0.1133	R	0.30575	12.00	3.669
72.00	10	0.0845	R	0.0989	24.00	2.3736
96.00	10	0	R	0.04225	24.00	1.014
						159.9659

Actual Time	**Animal**	**Conc**	**Treatment**	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	11	0	T	NA	NA	NA
0.50	11	3.3788	T	1.6894	0.50	0.8447
1.00	11	8.0922	T	5.7355	0.50	2.86775
2.00	11	9.5904	T	8.8413	1.00	8.8413
4.00	11	7.941	T	8.7657	2.00	17.5314
6.00	11	6.3055	T	7.12325	2.00	14.2465
9.00	11	4.075	T	5.19025	3.00	15.57075
12.00	11	2.6357	T	3.35535	3.00	10.06605
24.00	11	1.1382	T	1.88695	12.00	22.6434
36.00	11	0.6876	T	0.9129	12.00	10.9548
48.00	11	0.3857	T	0.53665	12.00	6.4398
72.00	11	0.1943	T	0.29	24.00	6.96
96.00	11	0	T	0.09715	24.00	2.3316
						119.2981

Actual Time	**Animal**	**Conc**	**Treatment**	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	11	0	R	NA	NA	NA
0.50	11	6.902	R	3.451	0.50	1.7255
1.00	11	8.0715	R	7.48675	0.50	3.743375
2.00	11	7.5862	R	7.82885	1.00	7.82885
4.00	11	6.3671	R	6.97665	2.00	13.9533
6.00	11	5.1882	R	5.77765	2.00	11.5553
9.00	11	3.5803	R	4.38425	3.00	13.15275
12.00	11	2.5373	R	3.0588	3.00	9.1764
24.00	11	2.1088	R	2.32305	12.00	27.8766
36.00	11	0.8041	R	1.45645	12.00	17.4774
48.00	11	0.3844	R	0.59425	12.00	7.131
72.00	11	0.0376	R	0.211	24.00	5.064
96.00	11	0	R	0.0188	24.00	0.4512
						119.1357

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}		Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	12	0	T	NA	NA	NA		0.00	12	0	R	NA	NA	NA
0.50	12	7.8037	T	3.90185	0.50	1.950925		0.50	12	14.7725	R	7.38625	0.50	3.693125
1.00	12	8.4611	T	8.1324	0.50	4.0662		1.00	12	13.0554	R	13.91395	0.50	6.956975
2.00	12	7.9048	T	8.18295	1.00	8.18295		2.00	12	11.8779	R	12.46665	1.00	12.46665
4.00	12	7.779	T	7.8419	2.00	15.6838		4.00	12	9.5409	R	10.7094	2.00	21.4188
6.00	12	6.3584	T	7.0687	2.00	14.1374		6.00	12	7.8499	R	8.6954	2.00	17.3908
9.00	12	4.6266	T	5.4925	3.00	16.4775		9.00	12	4.7223	R	6.2861	3.00	18.8583
12.00	12	3.2415	T	3.93405	3.00	11.80215		12.00	12	3.6077	R	4.165	3.00	12.495
24.00	12	1.4207	T	2.3311	12.00	27.9732		24.00	12	1.5017	R	2.5547	12.00	30.6564
36.00	12	1.2319	T	1.3263	12.00	15.9156		36.00	12	1.0317	R	1.2667	12.00	15.2004
48.00	12	0.5042	T	0.86805	12.00	10.4166		48.00	12	0.2023	R	0.617	12.00	7.404
72.00	12	0.215	T	0.3596	24.00	8.6304		72.00	12	0.1695	R	0.1859	24.00	4.4616
96.02	12	0.0972	T	0.1561	24.02	3.749002		96.00	12	0	R	0.08475	24.00	2.034
						138.9857								153.0361

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	13	0	T	NA	NA	NA
0.50	13	11.5164	T	5.7582	0.50	2.8791
1.00	13	12.5293	T	12.02285	0.50	6.011425
2.00	13	12.8171	T	12.6732	1.00	12.6732
4.00	13	12.6859	T	12.7515	2.00	25.503
6.00	13	11.1342	T	11.91005	2.00	23.8201
9.00	13	9.0335	T	10.08385	3.00	30.25155
12.00	13	7.864	T	8.44875	3.00	25.34625
24.00	13	1.6182	T	4.7411	12.00	56.8932
36.00	13	1.1014	T	1.3598	12.00	16.3176
48.00	13	0.2994	T	0.7004	12.00	8.4048
72.00	13	0.0992	T	0.1993	24.00	4.7832
96.00	13	0	T	0.0496	24.00	1.1904
						214.0738

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	13	0	R	NA	NA	NA
0.50	13	8.6753	R	4.33765	0.50	2.168825
1.00	13	10.2324	R	9.45385	0.50	4.726925
2.00	13	9.9372	R	10.0848	1.00	10.0848
4.00	13	9.806	R	9.8716	2.00	19.7432
6.00	13	8.2917	R	9.04885	2.00	18.0977
9.00	13	7.1497	R	7.7207	3.00	23.1621
12.00	13	5.3897	R	6.2697	3.00	18.8091
24.00	13	2.5165	R	3.9531	12.00	47.4372
36.00	13	1.2794	R	1.89795	12.00	22.7754
48.00	13	0.7405	R	1.00995	12.00	12.1194
72.00	13	0.2485	R	0.4945	24.00	11.868
96.00	13	0.0985	R	0.1735	24.00	4.164
						195.1567

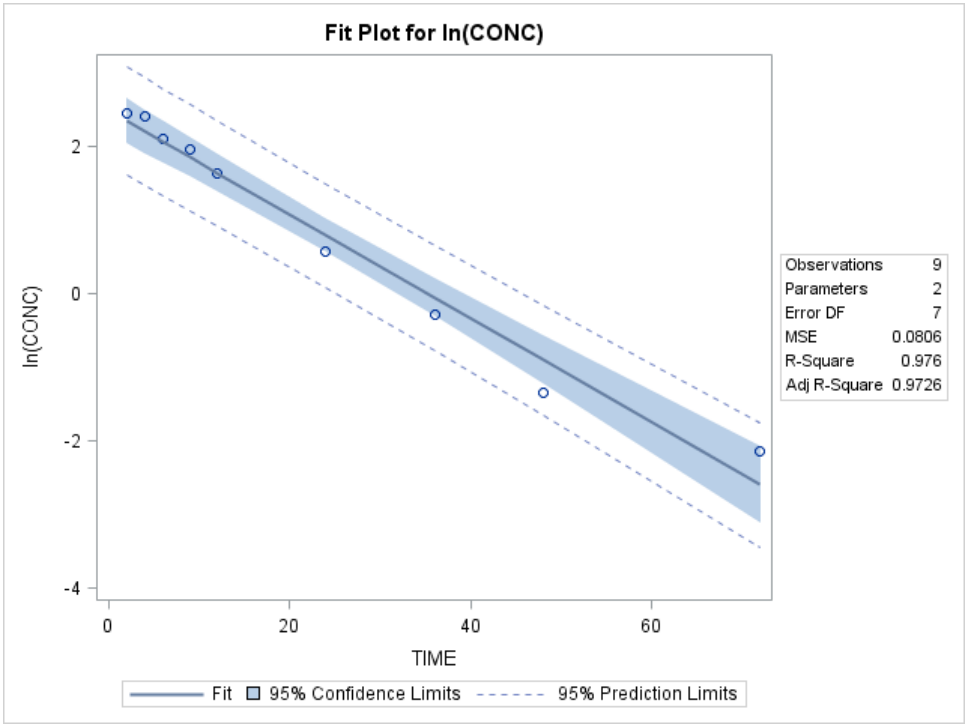
Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	14	0	T	NA	NA	NA
0.50	14	7.1283	T	3.56415	0.50	1.782075
1.00	14	11.4642	T	9.29625	0.50	4.648125
2.00	14	10.381	T	10.9226	1.00	10.9226
4.00	14	10.1862	T	10.2836	2.00	20.5672
6.00	14	8.6332	T	9.4097	2.00	18.8194
8.98	14	6.6623	T	7.64775	2.98	22.81579
12.00	14	5.286	T	5.97415	3.02	18.02202
24.00	14	2.7502	T	4.0181	12.00	48.2172
36.00	14	1.4059	T	2.07805	12.00	24.9366
48.00	14	0.5451	T	0.9755	12.00	11.706
72.00	14	0.0664	T	0.30575	24.00	7.338
96.00	14	0	T	0.0332	24.00	0.7968
						190.5718

Actual Time	Animal	Conc	Treatment	$\frac{C_i+C_{i-1}}{2}$	$T_i - T_{i-1}$	AUC_{0-t}
0.00	14	0	R	NA	NA	NA
0.50	14	15.699	R	7.8495	0.50	3.92475
1.00	14	15.6876	R	15.6933	0.50	7.84665
2.00	14	15.0303	R	15.35895	1.00	15.35895
4.00	14	13.3239	R	14.1771	2.00	28.3542
6.00	14	9.0957	R	11.2098	2.00	22.4196
9.00	14	7.7843	R	8.44	3.00	25.32
12.00	14	6.4883	R	7.1363	3.00	21.4089
24.00	14	2.6646	R	4.57645	12.00	54.9174
36.00	14	1.0833	R	1.87395	12.00	22.4874
48.00	14	0.4232	R	0.75325	12.00	9.039
72.00	14	0.1949	R	0.30905	24.00	7.4172
96.00	14	0	R	0.09745	24.00	2.3388
						220.8329

Appendix 3 Slope and Intercept (vertical axis) of figures in Appendix 4

Figure	Intercept (vertical intercept)	Slope
1	2.49478	-0.07069
2	2.59006	-0.08482
3	2.37959	-0.06371
4	2.13646	-0.05206
5	2.54009	-0.06666
6	2.58152	-0.07249
7	2.33197	-0.06469
8	2.72571	-0.08352
9	1.58798	-0.04670
10	2.47347	-0.08378
11	2.30245	-0.05683
12	3.08311	-0.10510
13	2.41261	-0.05375
14	3.39896	-0.11147
15	2.51812	-0.08338
16	2.93585	-0.09713
17	2.56782	-0.06314
18	1.42249	-0.03254
19	2.45498	-0.07787
20	2.88456	-0.09655
21	0.94957	-0.03686
22	2.11897	-0.07047
23	1.51104	-0.04126
24	2.39776	-0.06724
25	2.74958	-0.07474
26	2.33186	-0.05103
27	2.57993	-0.06986
28	2.68121	-0.06572

Appendix 4 Finding λ on Calculating $AUC_{0-\infty}$



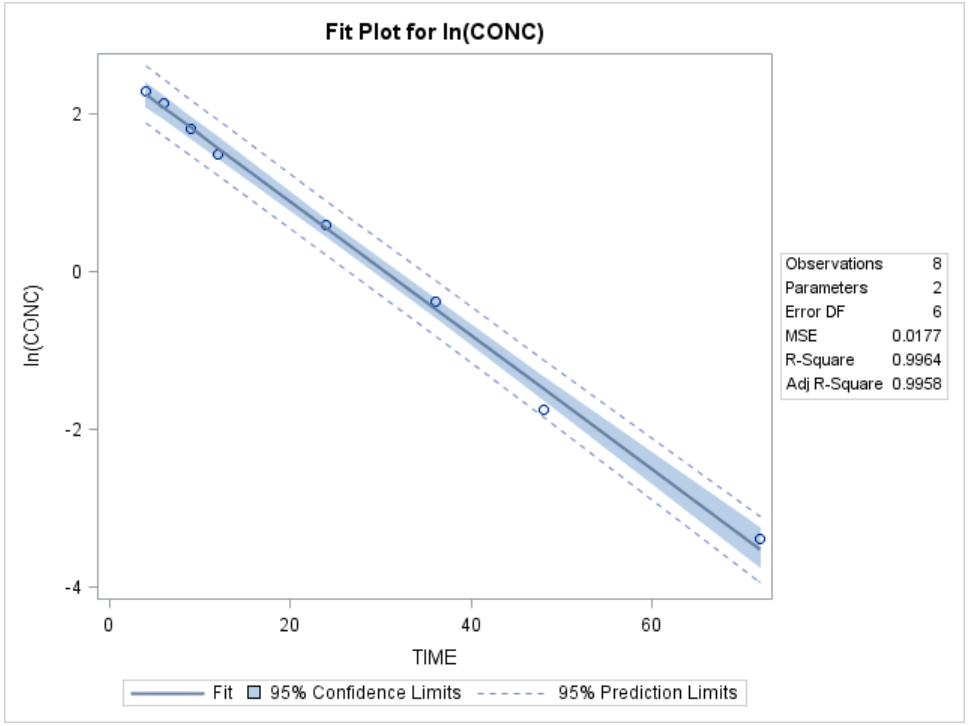
C_t = 0.1173

λ = (-0.07069)*(-2.303)

= 0.162799

Figure 1: Fit Plot for In(Conc) for Sheep 1,

Test Treatment in Period 2



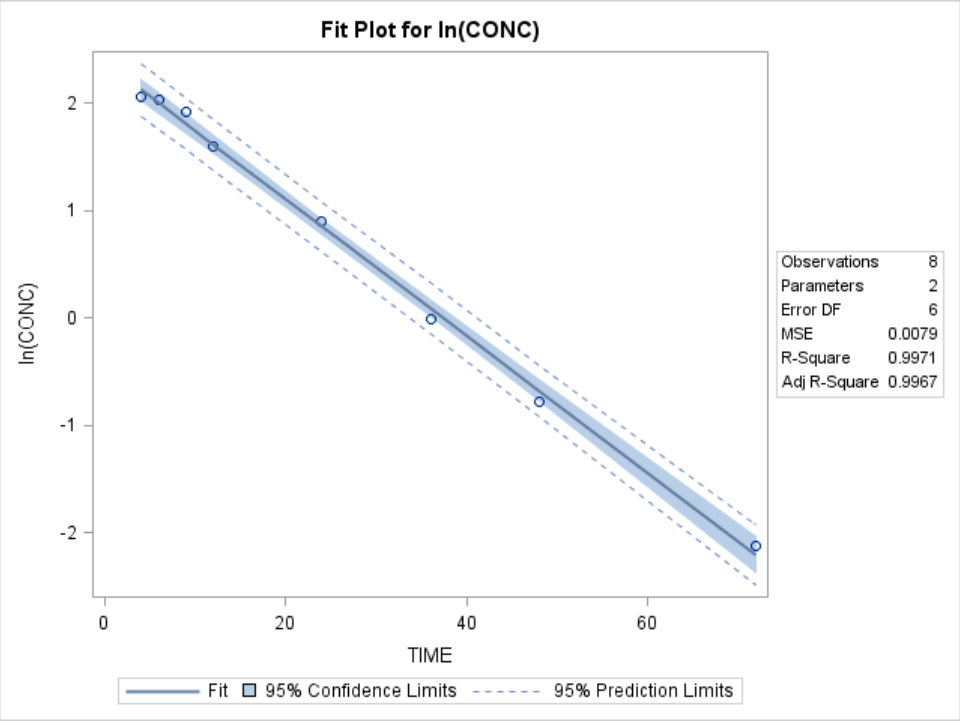
C_t = 0.0336

λ = (-0.08482)*(-2.303)

= 0.19534

Figure 2: Fit Plot for In(Conc) for Sheep 1,

Reference Treatment in Period 1

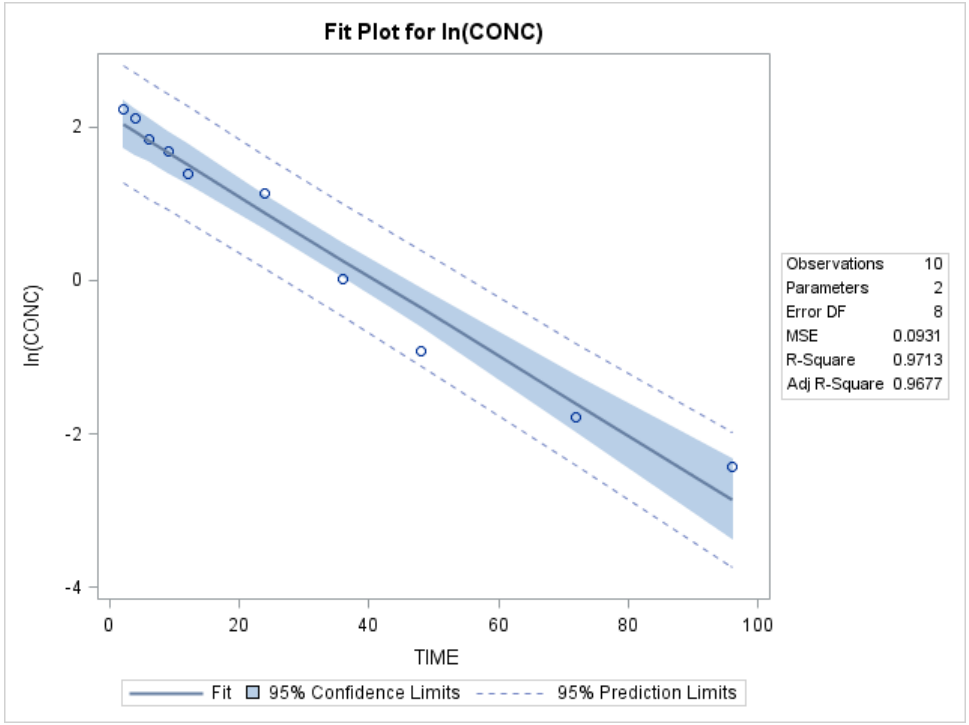


C_t = 0.1206

λ = (-0.06371)*(-2.303)

= 0.146724

Figure 3: Fit Plot for In(Conc) for Sheep 2,
Reference Treatment in Period 1

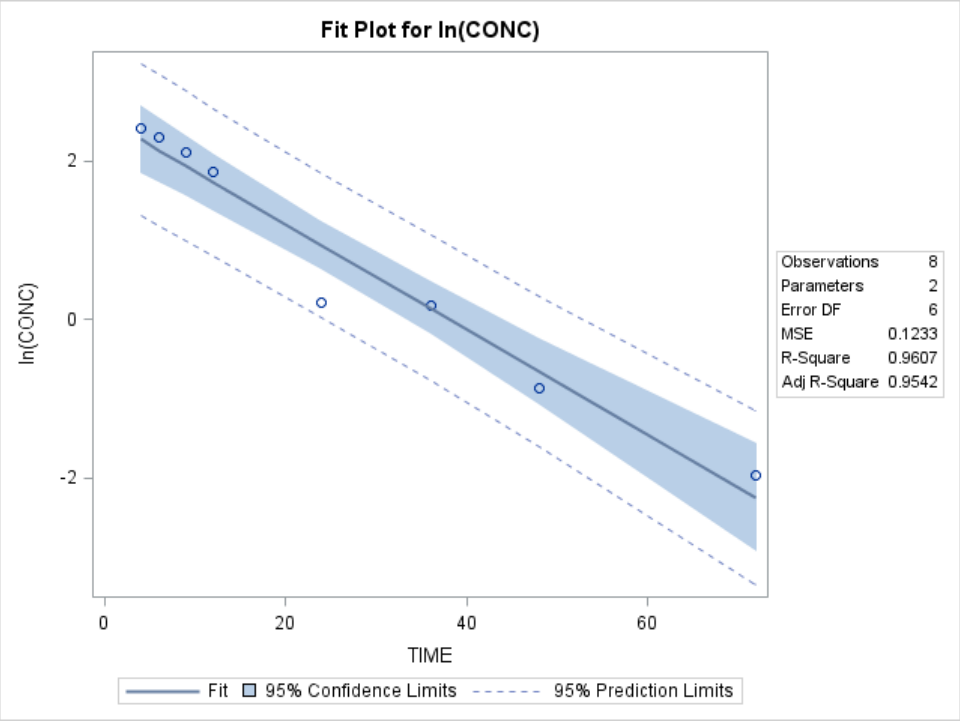


C_t = 0.0885

λ = (-0.05206)*(-2.303)

= 0.119894

Figure 4: Fit Plot for In(Conc) for Sheep 2,
Test Treatment in Period 2

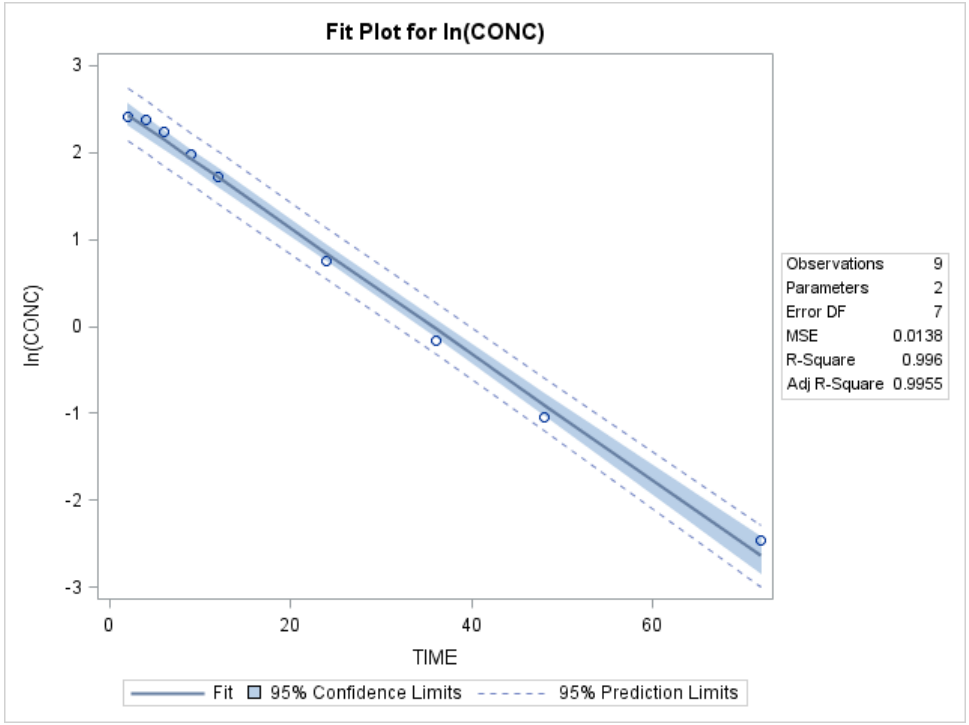


C_t = 0.1407

λ = (-0.06666)*(-2.303)

= 0.15338

Figure 5: Fit Plot for In(Conc) for Sheep 3,
Test Treatment in Period 2

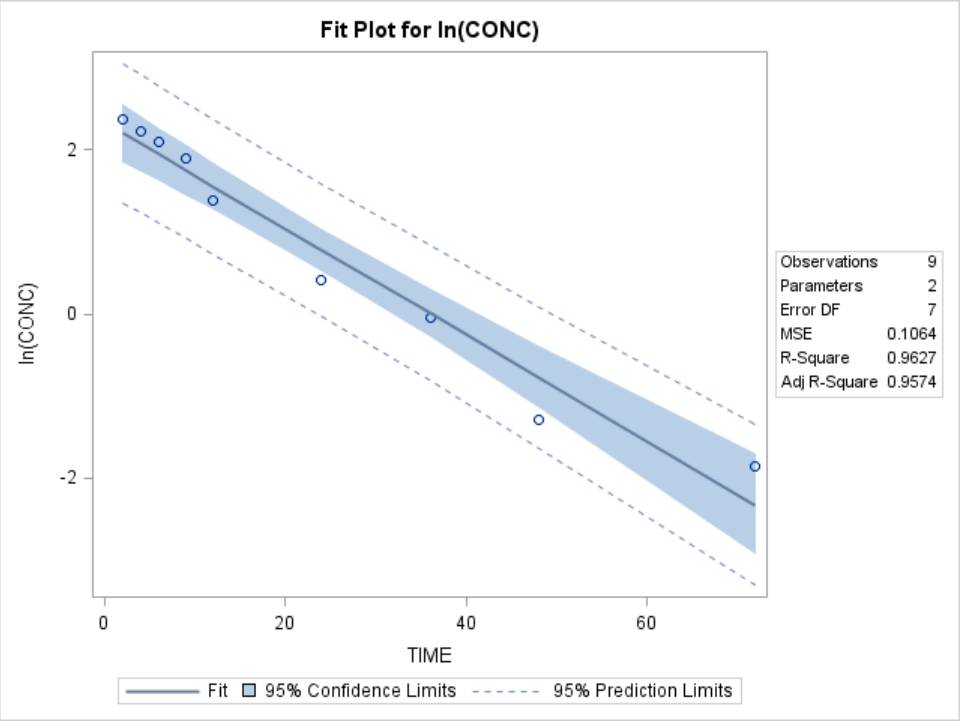


C_t=0.0852

λ =(-0.07249)*(-2.0303)

= 0.166944

Figure 6: Fit Plot for In(Conc) for Sheep 3,
Reference Treatment in Period 1

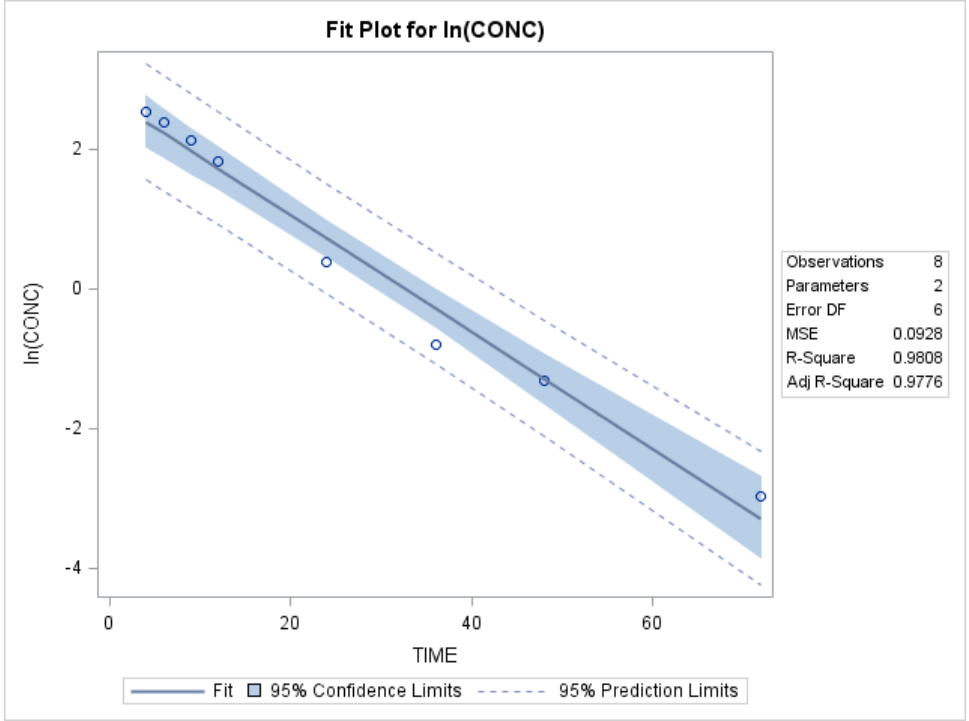


C_t = 0.1561

λ = (-0.06469)*(-2.303)

= 0.148981

Figure 7: Fit Plot for In(Conc) for Sheep 4,
Reference Treatment in Period 1

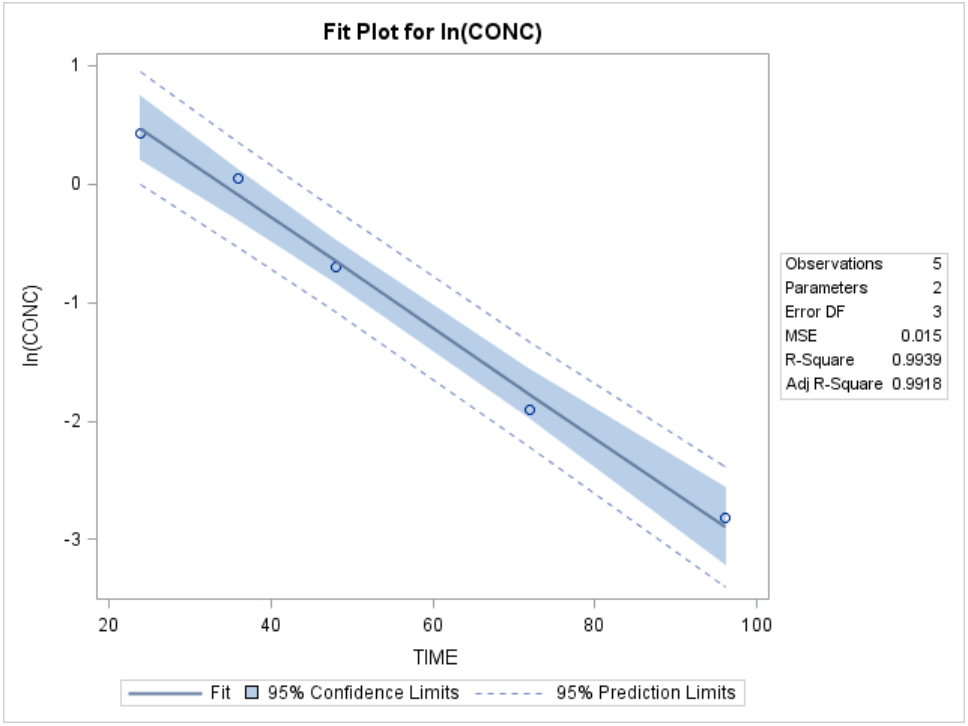


C_t = 0.0517

λ = (-0.08352)*(-2.303)

= 0.192347

Figure 8: Fit Plot for In(Conc) for Sheep 4,
Test Treatment in Period 2

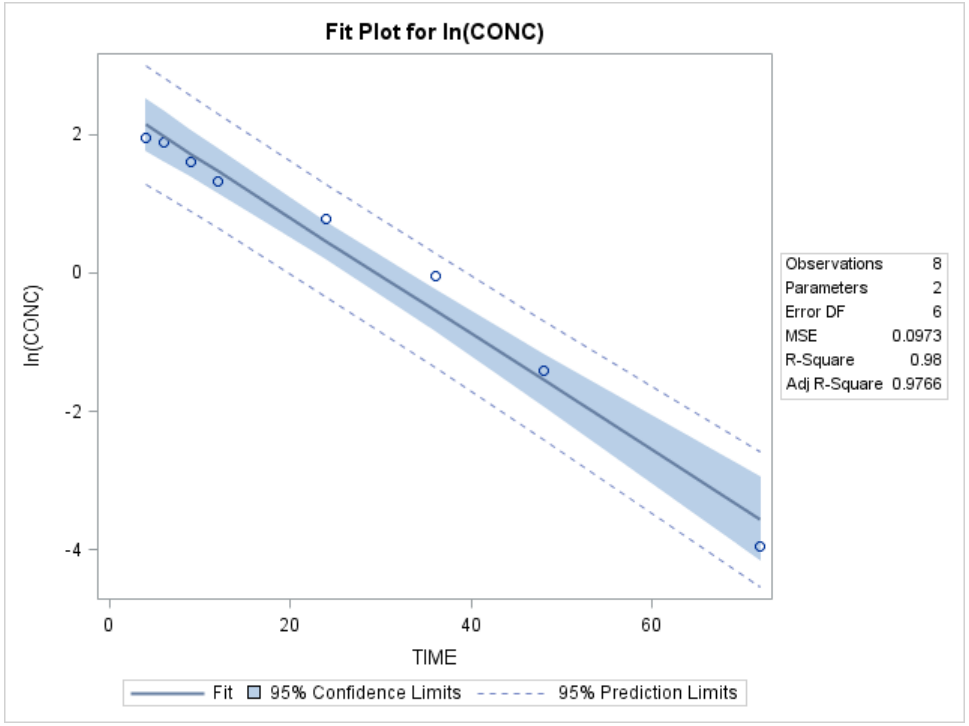


C_t = 0.0597

λ = (-0.04670)*(-2.303)

= 0.10755

Figure 9: Fit Plot for In(Conc) for Sheep 5,
Reference Treatment in Period 2

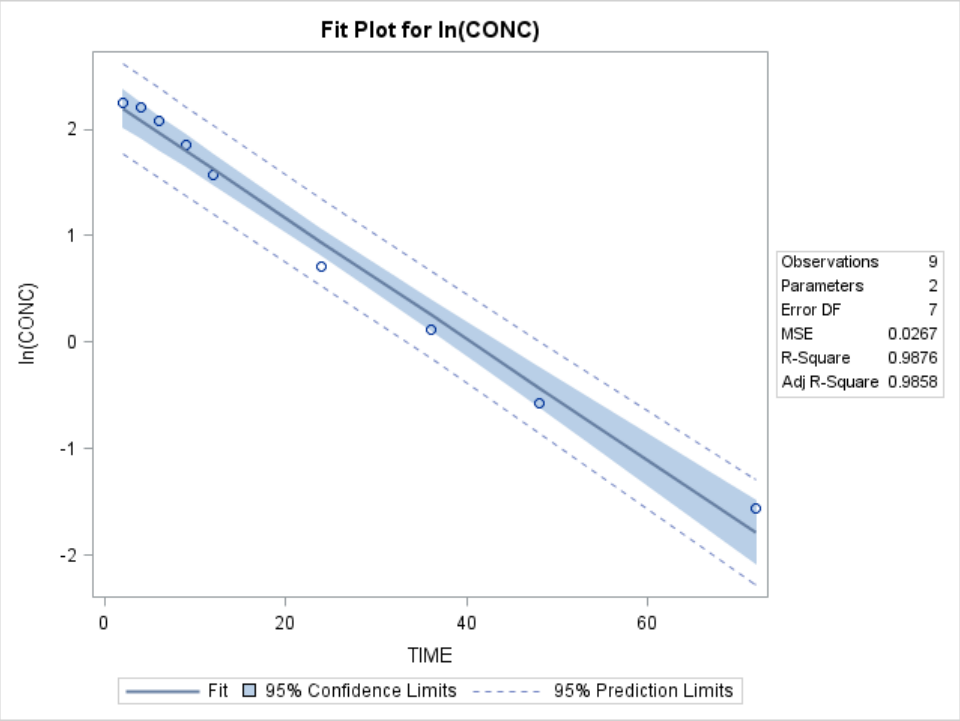


C_t = 0.0195

λ = (-0.08378)*(-2.303)

= 0.192945

Figure 10: Fit Plot for In(Conc) for Sheep 5,
Test Treatment in Period 1

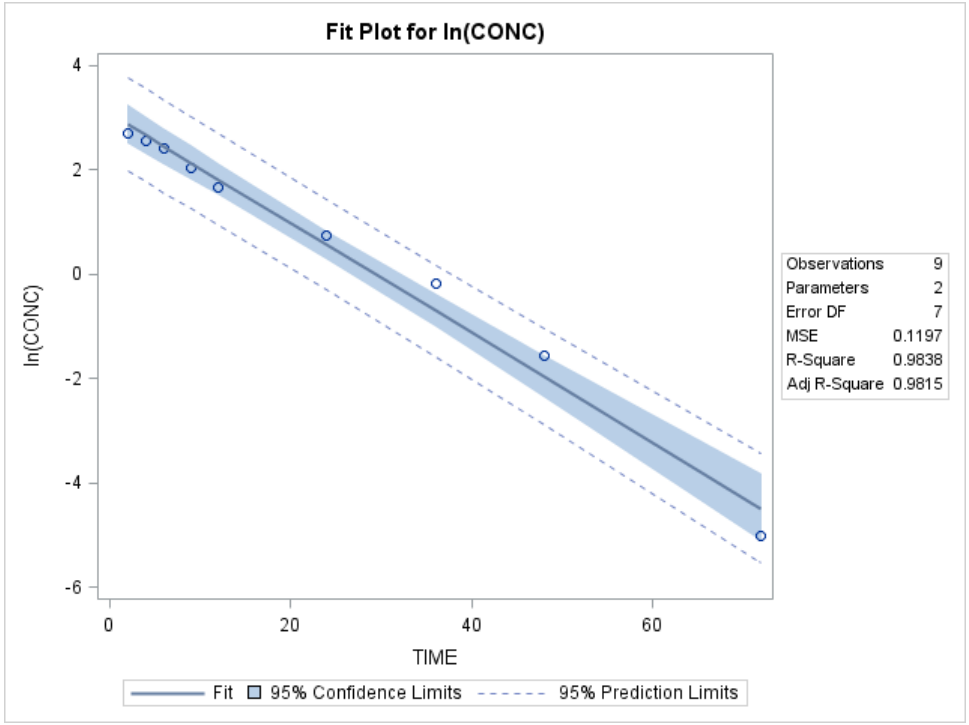


C_t = 0.2103

λ = (-0.05683)*(-2.303)

= 0.130879

Figure 11: Fit Plot for In(Conc) for Sheep 6,
Test Treatment in Period 1

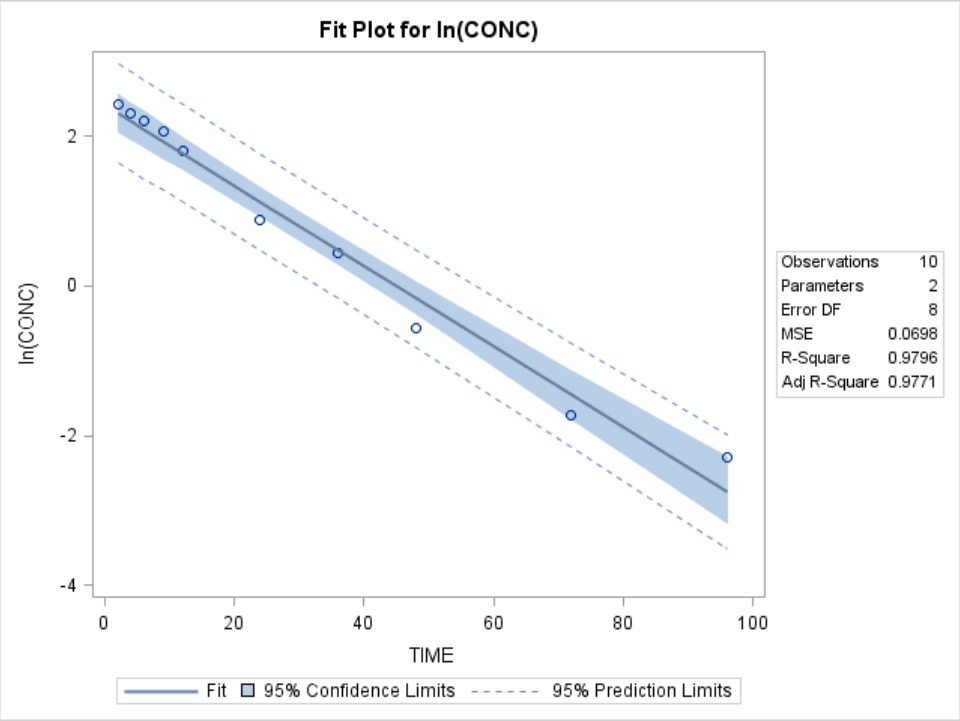


C_t = 0.0066

λ = (-0.1051)*(-2.303)

= 0.242045

Figure 12: Fit Plot for In(Conc) for Sheep 6,
Reference Treatment in Period 2

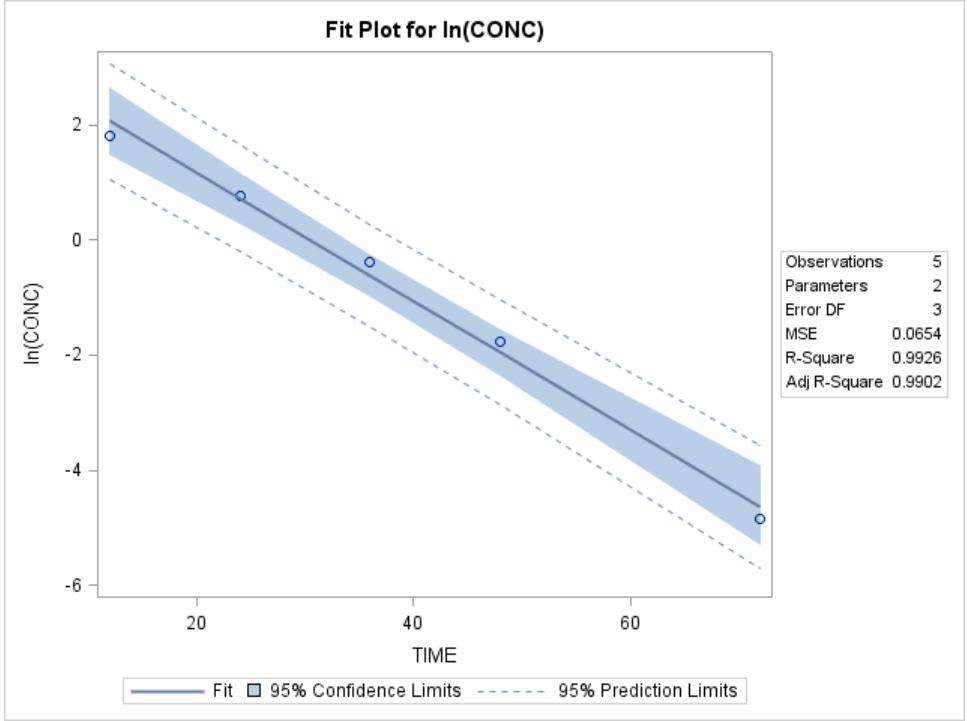


C_t = 0.1005

λ = (-0.05375)*(-2.303)

= 0.123786

Figure 13: Fit Plot for In(Conc) for Sheep 7,
Test Period in Period 2

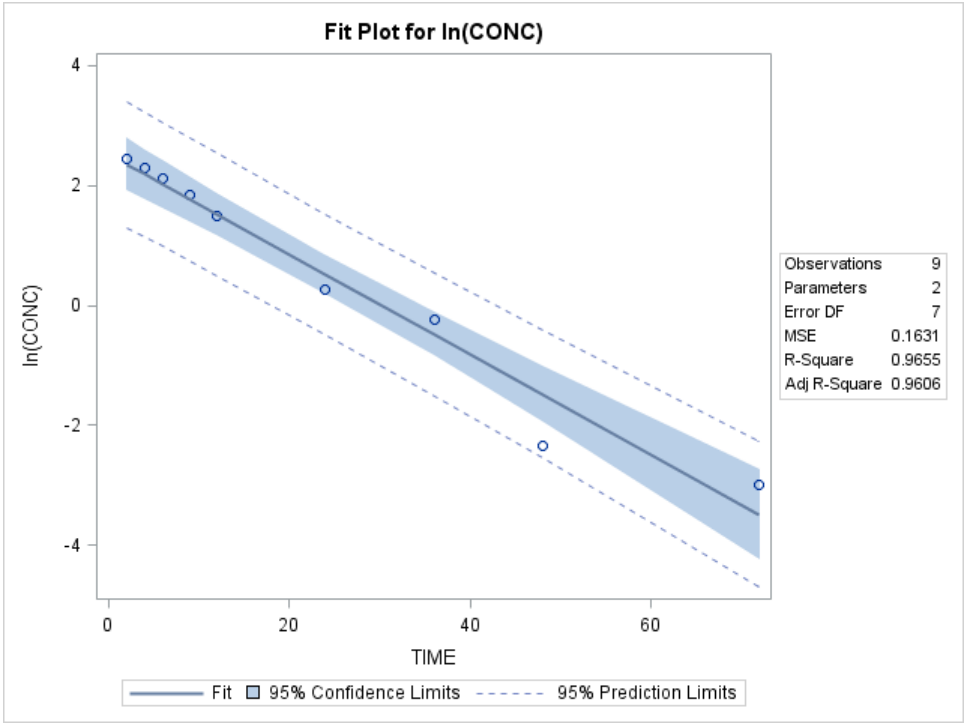


C_t = 0.0079

λ = (-0.11147)*(-2.303)

= 0.256715

Figure 14: Fit Plot for In(Conc) for Sheep 7,
Reference Treatment in Period 1

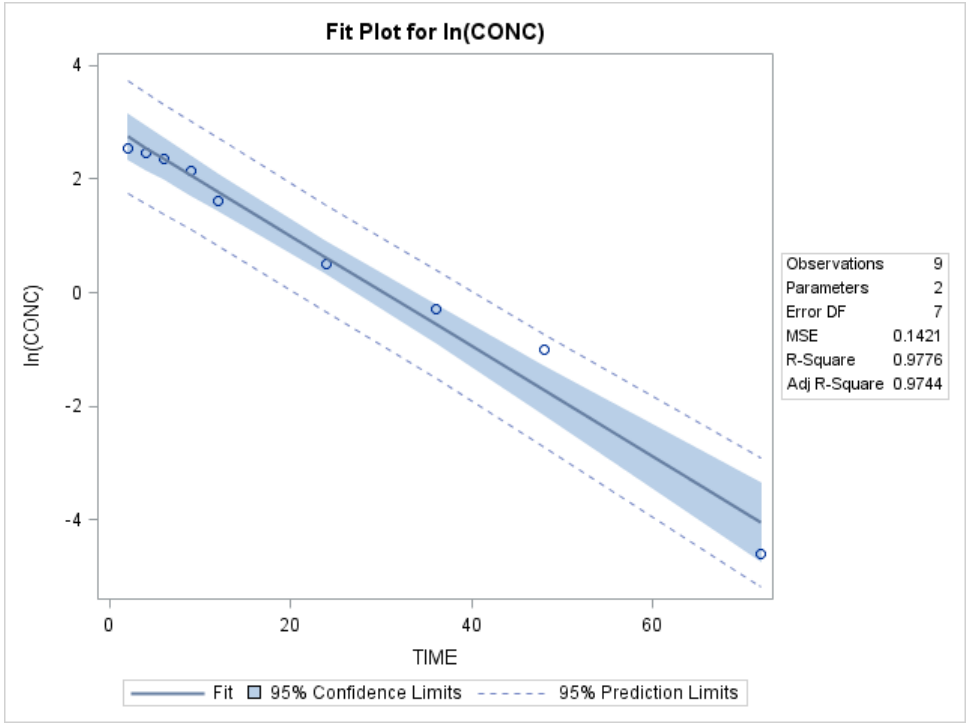


C_t = 0.0507

λ = (-0.08338)*(-2.303)

= 0.192024

Figure 15: Fit Plot for In(Conc) for Sheep 8,
Reference Treatment in Period 1

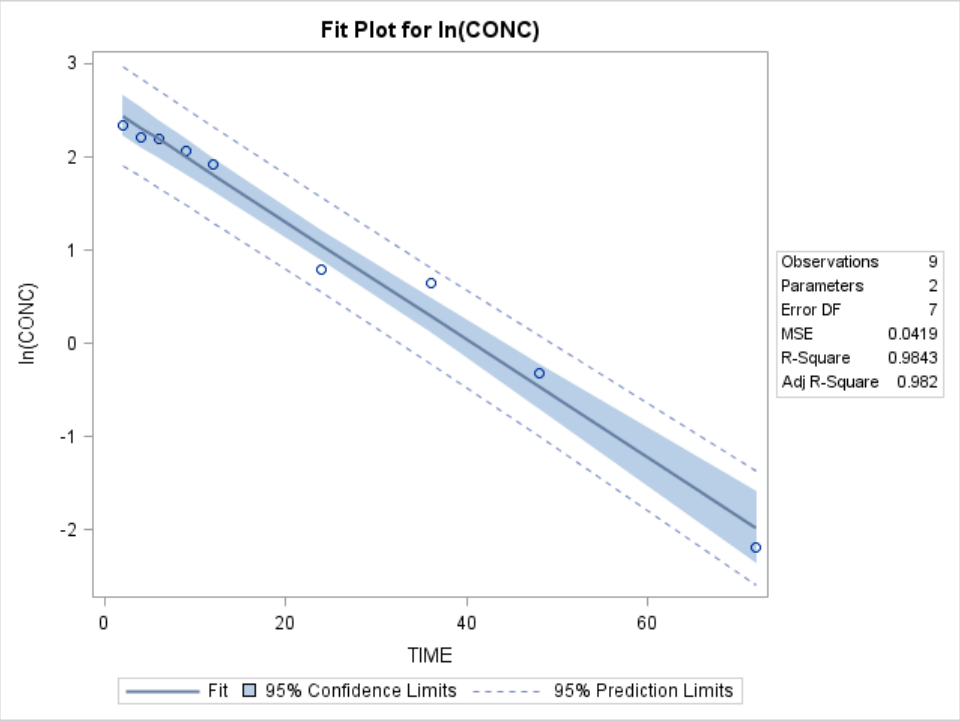


C_t = 0.0099

λ = (-0.09713)*(-2.303)

= 0.22369

Figure 16: Fit Plot for In(Conc) for Sheep 8,
Test Treatment in Period 2

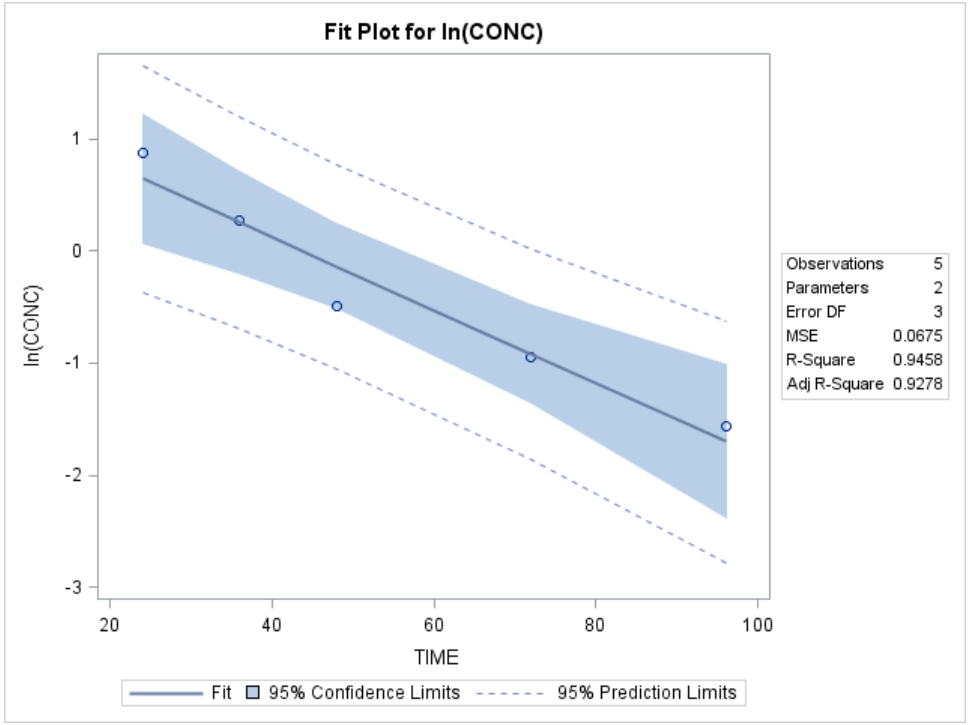


C_t = 0.1126

λ = (-0.06314)*(-2.303)

= 0.145411

Figure 17: Fit Plot for In(Conc) for Sheep 9,
Reference Treatment in Period 2

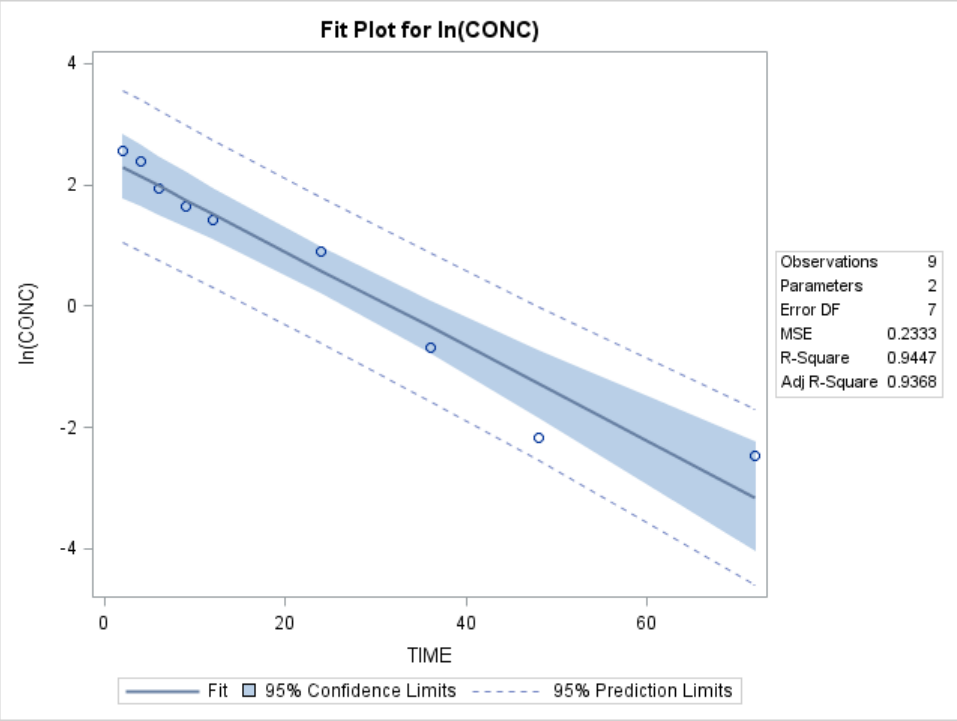


C_t = 0.209

λ = (-0.03254)*(-2.303)

=0.07494

Figure 18: Fit Plot for In(Conc) for Sheep 9,
Test Treatment in Period 1

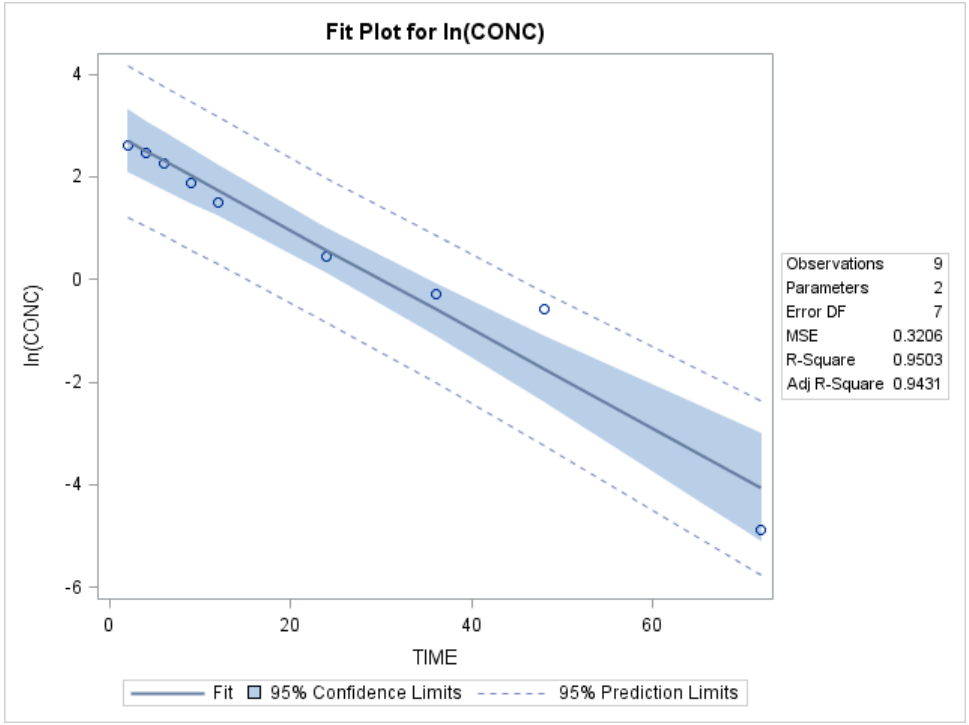


C_t = 0.0845

λ = (-0.07787)*(-2.303)

= 0.179335

Figure 19: Fit Plot for In(Conc) for Sheep 10,
Reference Treatment in Period 2

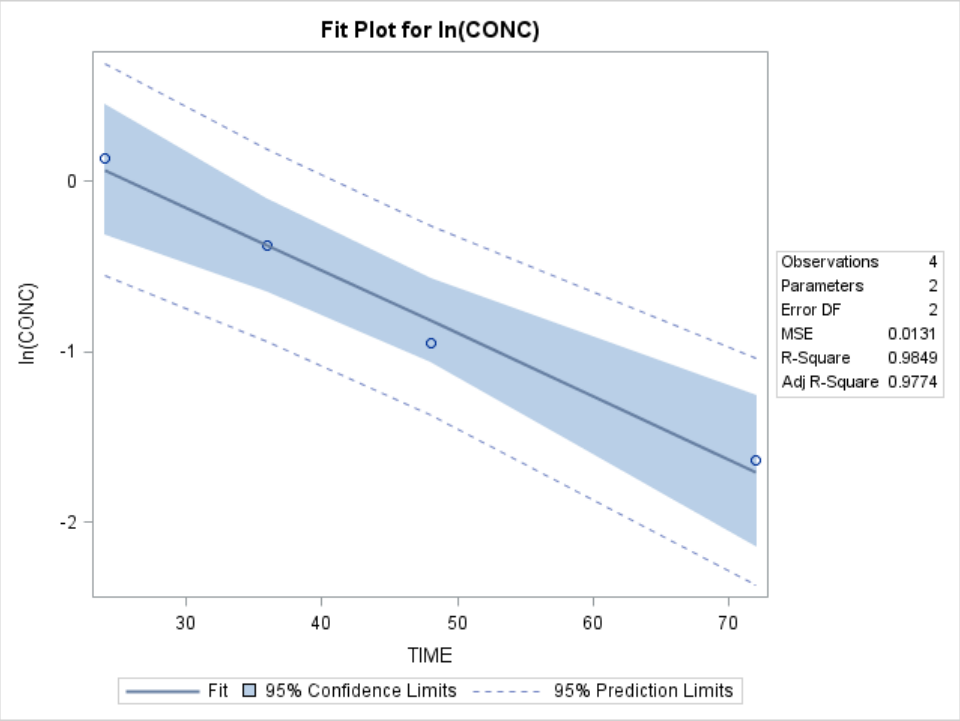


C_t = 0.0075

λ = (-0.09655)*(-2.303)

= 0.222355

Figure 20: Fit Plot for In(Conc) for Sheep 10,
Test Treatment in Period 1

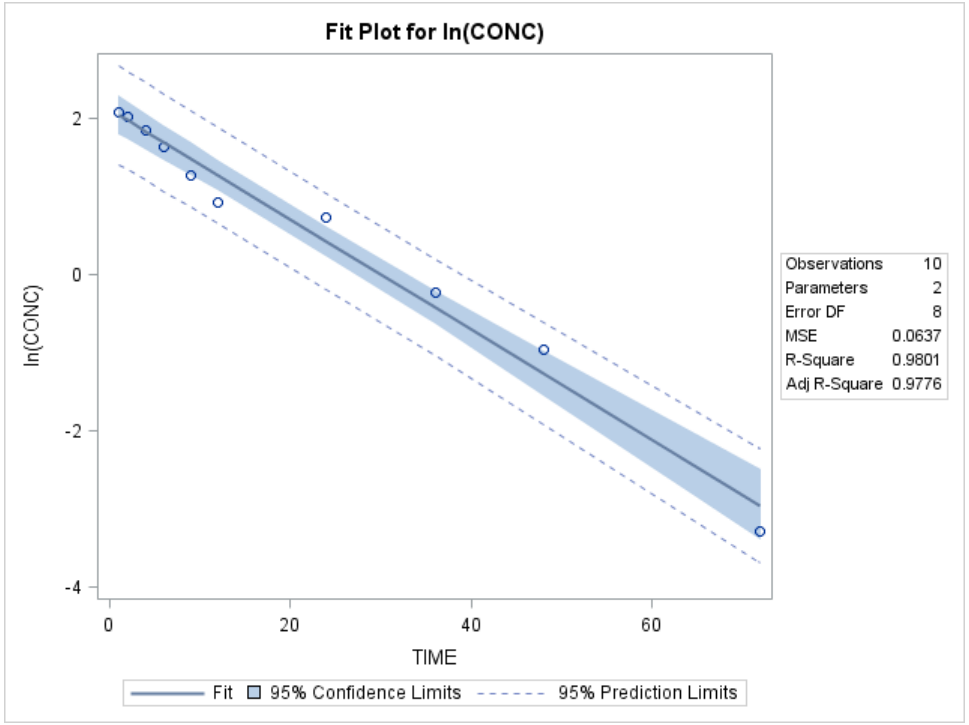


C_t = 0.1943

λ = (-0.03686)*(-2.303)

= 0.084889

Figure 21: Fit Plot for In(Conc) for Sheep 11,
Test Treatment in Period 1

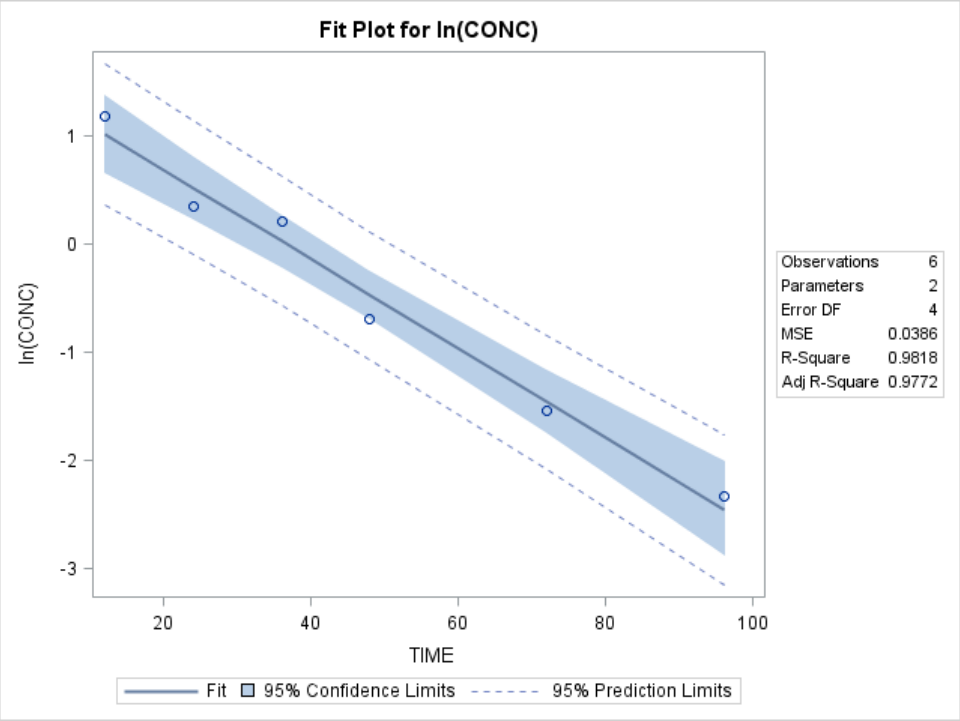


C_t = 0.0376

λ = (-0.07047)*(-2.303)

= 0.162292

Figure 22: Fit Plot for In(Conc) for Sheep 11,
Reference Treatment in Period 2

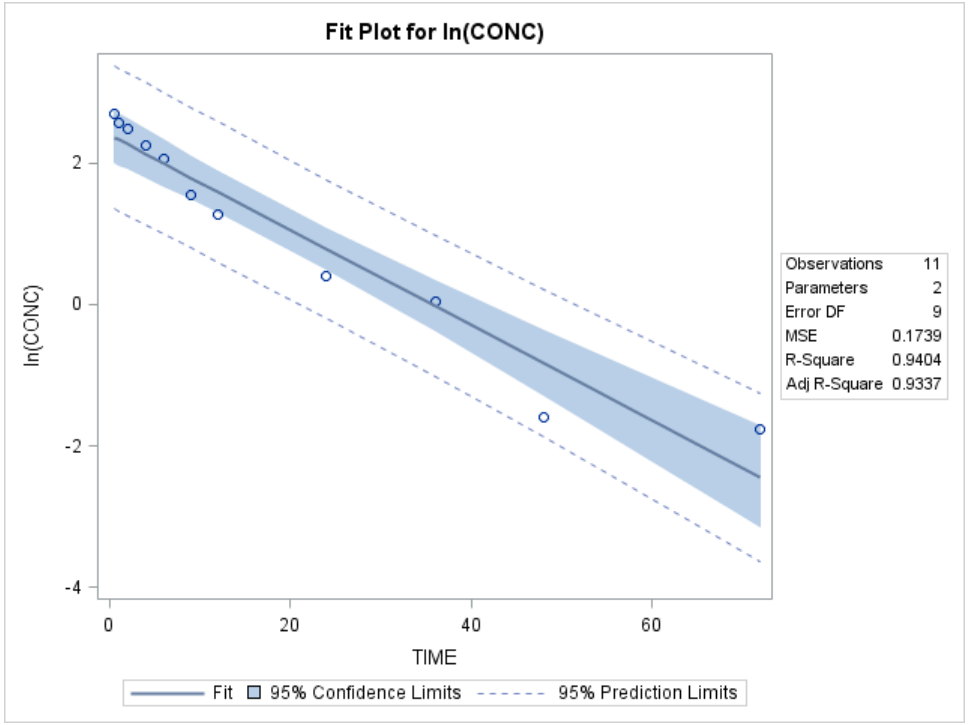


C_t = 0.0972

λ = (-0.04126)*(-2.303)

= 0.095022

Figure 23: Fit Plot for In(Conc) for Sheep 12,
Test Treatment in Period 1

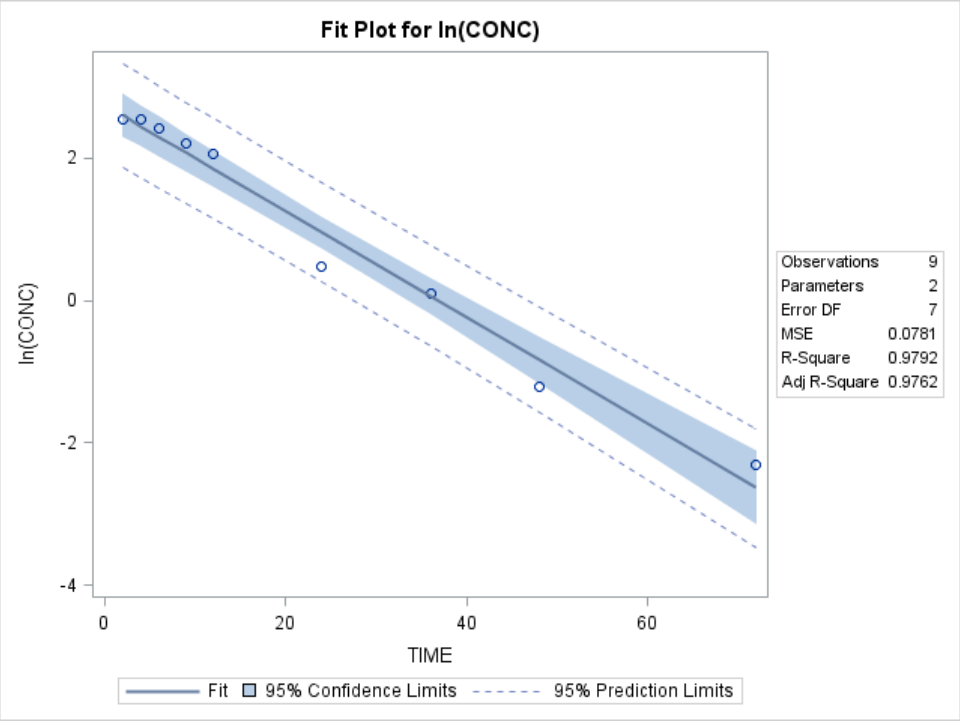


C_t = 0.1695

λ = (-0.06724)*(-2.303)

= 0.154854

Figure 24: Fit Plot for In(Conc) for Sheep 12,
Reference Treatment in Period 2

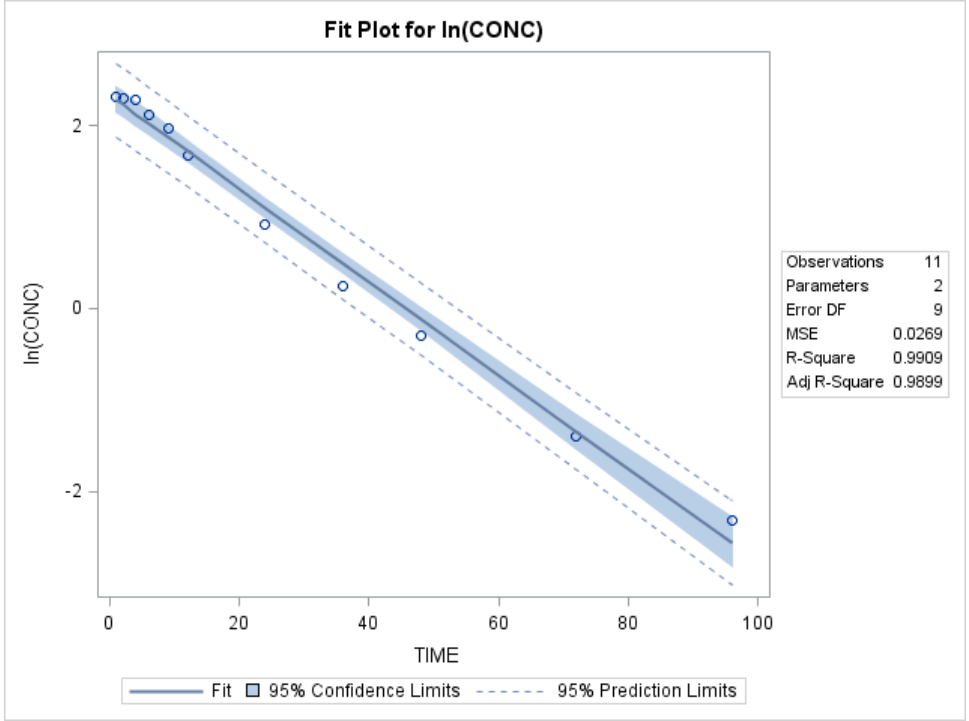


C_t = 0.0992

λ = (-0.07474)*(-2.303)

= 0.172126

Figure 25: Fit Plot for In(Conc) for Sheep 13,
Test Treatment in Period 2

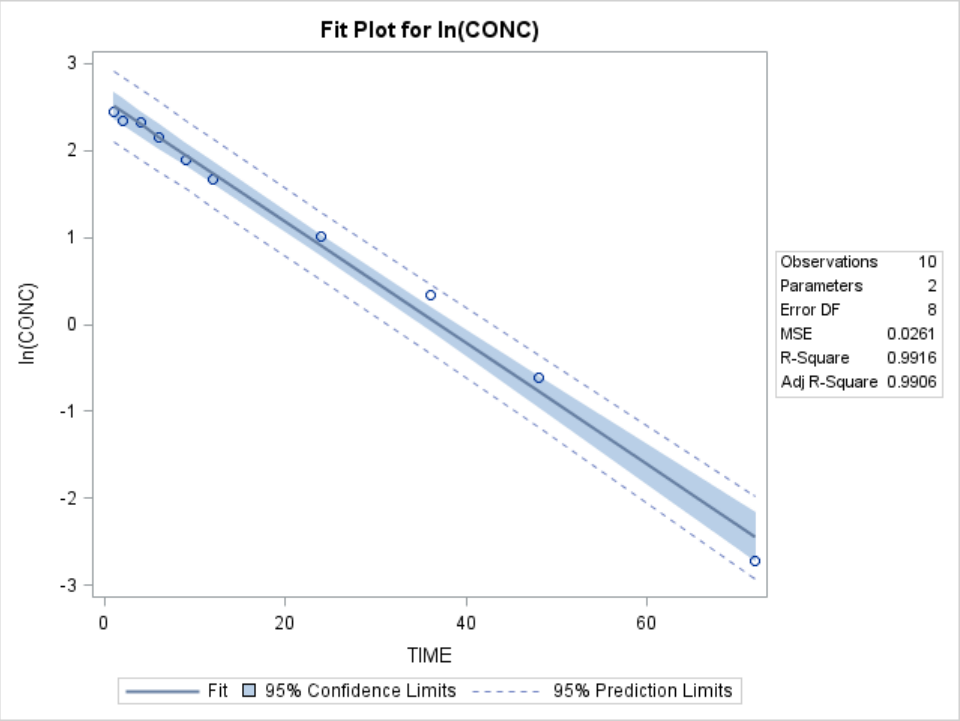


C_t = 0.0985

λ = (-0.05103)*(-2.303)

= 0.117522

Figure 26: Fit Plot for In(Conc) for Sheep 13,
Reference Treatment in Period 1

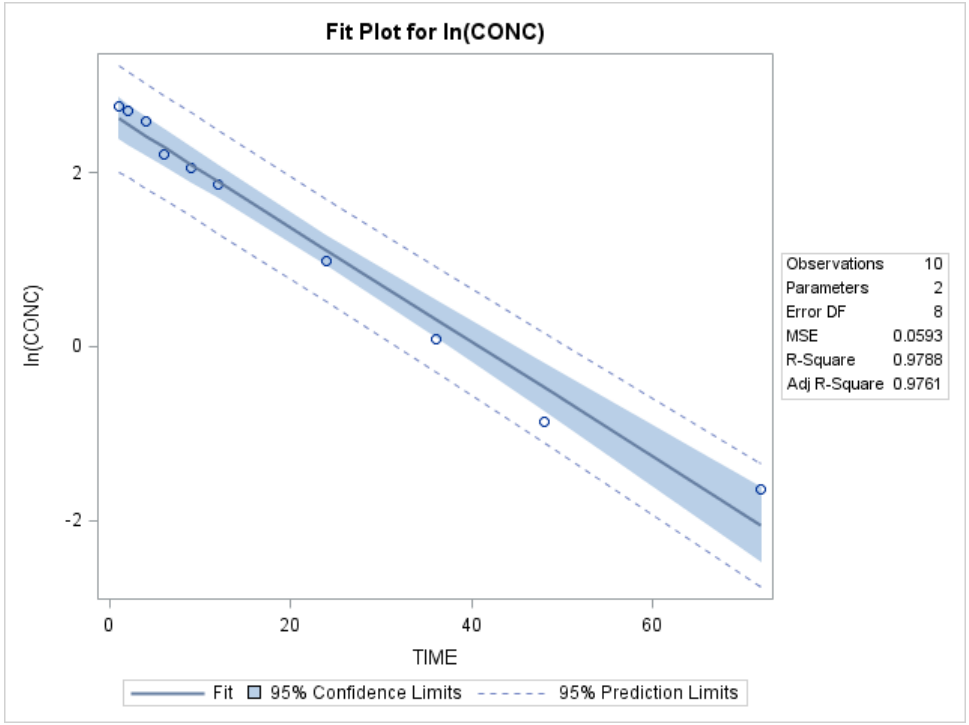


C_t = 0.0664

λ = (-0.06986)*(-2.303)

= 0.160888

Figure 27: Fit Plot for In(Conc) for Sheep 14,
Test Treatment in Period 1



C_t = 0.1949

λ = (-0.06572)*(-2.303)

= 0.151353

Figure 28: Fit Plot for In(Conc) for Sheep 14,
Reference Treatment in Period 2

Appendix 5 Calculating *AUC* from time zero to infinity

Sequence 1

	Animal	AUC_{0-t}	C_t	Slope	Constant	ƛ	$AUC_{0-\infty}$	$lnAUC_{0-\infty}$
Test	5	132.0207	0.0195	-0.08378	-2.303	0.192945	132.1218	4.883724
	6	170.1138	0.2103	-0.05683	-2.303	0.130879	171.7206	5.145869
	9	222.0024	0.209	-0.03254	-2.303	0.07494	224.7913	5.415172
	10	173.6732	0.0075	-0.09655	-2.303	0.222355	173.7069	5.15737
	11	119.2981	0.1943	-0.03686	-2.303	0.084889	121.587	4.80063
	12	138.9857	0.0972	-0.04126	-2.303	0.095022	140.0086	4.941704
	14	190.5718	0.0664	-0.06986	-2.303	0.160888	190.9845	5.252192
Reference	5	170.5818	0.0597	-0.0467	-2.303	0.10755	171.1369	5.142464
	6	196.4537	0.0066	-0.1051	-2.303	0.242045	196.481	5.280566
	9	207.0228	0.1126	-0.06314	-2.303	0.145411	207.7972	5.336562
	10	159.9659	0.0845	-0.07787	-2.303	0.179335	160.4371	5.077902
	11	119.1357	0.0376	-0.07047	-2.303	0.162292	119.3674	4.782206
	12	153.0361	0.1695	-0.06724	-2.303	0.154854	154.1307	5.037801
	14	220.8329	0.1949	-0.06572	-2.303	0.151353	222.1206	5.403221

Sequence 2

	Animal	AUC_{0-t}	C_t	Slope	Constant	ƛ	$AUC_{0-\infty}$	
Reference1	1	167.1385	0.0336	-0.08482	-2.303	0.19534	167.3105	5.119851
	2	166.5281	0.1206	-0.06371	-2.303	0.146724	167.3501	5.120088
	3	181.6076	0.0852	-0.07249	-2.303	0.166944	182.1179	5.204655
	4	152.5701	0.1561	-0.06469	-2.303	0.148981	153.6179	5.034468
	7	174.1533	0.0079	-0.11147	-2.303	0.256715	174.1841	5.160113
	8	152.1165	0.0507	-0.08338	-2.303	0.192024	152.3805	5.026381
	13	195.1567	0.0985	-0.05103	-2.303	0.117522	195.9948	5.278088
Test	1	170.3735	0.1173	-0.07069	-2.303	0.162799	171.094	5.142213
	2	164.6649	0.0885	-0.05206	-2.303	0.119894	165.4031	5.108385
	3	187.0004	0.1407	-0.0666	-2.303	0.15338	187.9177	5.236004
	4	184.3159	0.0517	-0.08352	-2.303	0.192347	184.5847	5.218108
	7	201.7071	0.1005	-0.05375	-2.303	0.123786	202.519	5.310834
	8	183.1658	0.0099	-0.09713	-2.303	0.22369	183.2101	5.210633
	13	214.0738	0.0992	-0.07474	-2.303	0.172126	214.6501	5.369009

Appendix 6 Wilcoxon Mann-Whitney Two One-Sided Tests Calculations

Sequence	Subject	Period		Diff	L	Ranks L	U	Ranks U
1	Number	1	2					
RT	5	2	2	0	0.314	8	-0.314	3
RT	6	1	2	0.5	0.814	12.5	0.186	12.5
RT	9	2	9	3.5	3.814	14	3.186	14
RT	10	2	2.02	0.01	0.324	9	-0.304	4
RT	11	1	2	0.5	0.814	12.5	0.186	12.5
RT	12	0.5	1	0.25	0.564	10.5	-0.064	5.5
RT	14	0.5	1	0.25	0.564	10.5	-0.064	5.5
2								
TR	1	2	2	0	0	3.5	0	7.5
TR	2	2	2	0	0	3.5	0	7.5
TR	3	4	2	-1	-1	1	-1	1
TR	4	2	2	0	0	3.5	0	7.5
TR	7	2	2	0	0	3.5	0	7.5
TR	8	2	2	0	0	3.5	0	7.5
TR	13	2	1	-0.5	-0.5	2	-0.5	2

Appendix 7 Outliers Plot and Histogram for $AUC_{0-\infty}$ untransformed data.

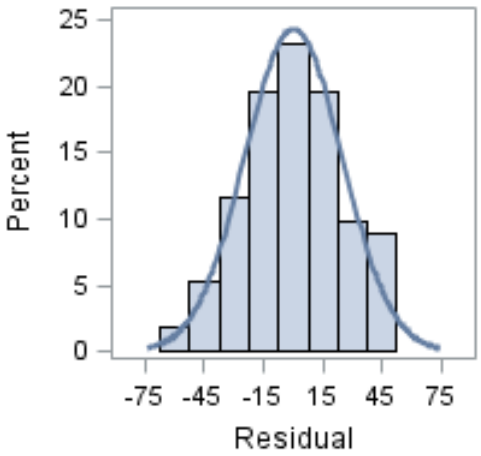


Figure 1 Fit Histogram for Residuals of $AUC_{0-\infty}$ untransformed data

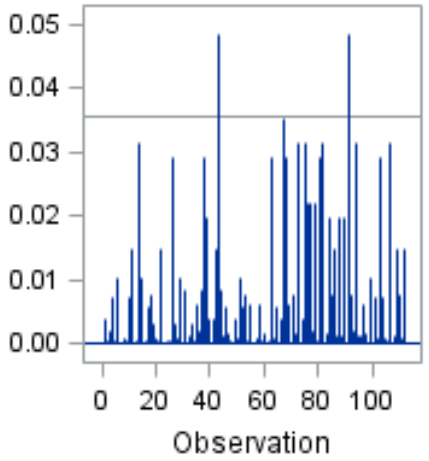


Figure 2 Fit Outliers Plot for $AUC_{0-\infty}$ untransformed data

Appendix 8 Outliers Plot and Histogram for C_{max} untransformed data.

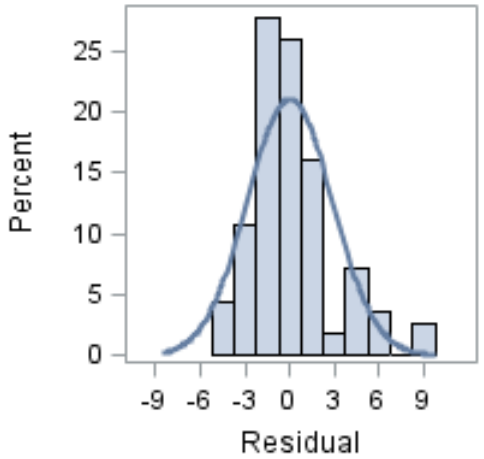


Figure 1 Fit Histogram for Residuals of C_{max} untransformed data

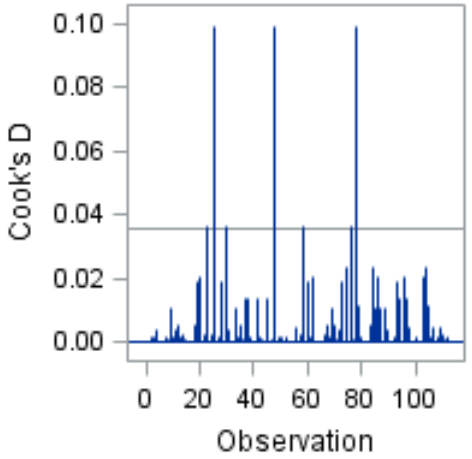


Figure 2 Fit Outliers Plot for C_{max} untransformed data

Appendix 9 Bootstrap Samples of $AUC_{0-\infty}$ and C_{max}

Bootstrap Samples of $AUC_{0-\infty}$	$lnAUC_{0-\infty}$		Bootstrap Samples of C_{max}	lnC_{max}
160.4371	5.077902		11.4669	2.439465
171.7206	5.145869		12.8158	2.550679
187.9177	5.236004		10.3462	2.336619
154.1307	5.037801		9.7471	2.27697
171.094	5.142213		11.4642	2.439229
202.519	5.310834		11.9468	2.480463
174.1841	5.160113		10.3462	2.336619
184.5847	5.218108		10.7164	2.371775
171.1369	5.142464		8.5327	2.143906
154.1307	5.037801		12.6993	2.541547
207.7972	5.336562		9.7471	2.27697
174.1841	5.160113		9.4104	2.241815
171.7206	5.145869		10.3462	2.336619
222.1206	5.403221		10.2324	2.325559
202.519	5.310834		10.7164	2.371775
171.094	5.142213		11.2346	2.418998
195.9948	5.278088		11.0619	2.403507
153.6179	5.034468		9.4104	2.241815
190.9845	5.252192		15.699	2.753597
184.5847	5.218108		7.3284	1.991757
171.1369	5.142464		11.4669	2.439465

Bootstrap Samples of $AUC_{0-\infty}$	$lnAUC_{0-\infty}$		Bootstrap Samples of C_{max}	lnC_{max}
207.7972	5.336562		10.2324	2.325559
173.7069	5.15737		17.303	2.85088
171.1369	5.142464		10.2324	2.325559
182.118	5.204655		21	3.044522
132.1218	4.883724		11.2346	2.418998
190.9845	5.252192		12.8171	2.55078
183.2101	5.210633		15.699	2.753597
202.519	5.310834		11.2346	2.418998
174.1841	5.160113		17.303	2.85088
152.3805	5.026381		9.7471	2.27697
167.3501	5.120088		11.2346	2.418998
190.9845	5.252192		8.4611	2.135479
182.118	5.204655		12.8171	2.55078
196.481	5.280566		9.4104	2.241815
187.9177	5.236004		10.7164	2.371775
152.3805	5.026381		8.0715	2.088339
132.1218	4.883724		8.0715	2.088339
140.0086	4.941704		12.7964	2.549164
160.4371	5.077902		11.4669	2.439465
160.4371	5.077902		8.0715	2.088339
207.7972	5.336562		12.7964	2.549164
119.3674	4.782206		10.7164	2.371775

Bootstrap Samples of $AUC_{0-\infty}$	$lnAUC_{0-\infty}$		Bootstrap Samples of C_{max}	lnC_{max}
152.3805	5.026381		11.0619	2.403507
167.3501	5.120088		8.0715	2.088339
195.9948	5.278088		11.4642	2.439229
165.4031	5.108385		10.7164	2.371775
171.094	5.142213		21	3.044522
160.4371	5.077902		12.8171	2.55078
184.5847	5.218108		12.8158	2.550679
202.519	5.310834		11.0619	2.403507
195.9948	5.278088		12.8158	2.550679
153.6179	5.034468		11.1924	2.415235
182.118	5.204655		11.9468	2.480463
196.481	5.280566		11.1924	2.415235
173.7069	5.15737		9.5904	2.260763
184.5847	5.218108		10.2324	2.325559
196.481	5.280566		17.303	2.85088
171.1369	5.142464		11.0619	2.403507
165.4031	5.108385		15.699	2.753597
174.1841	5.160113		12.7964	2.549164
171.094	5.142213		7.3284	1.991757
132.1218	4.883724		11.9468	2.480463
184.5847	5.218108		11.0619	2.403507
195.9948	5.278088		11.0619	2.403507

Bootstrap Samples of $AUC_{0-\infty}$	$lnAUC_{0-\infty}$		Bootstrap Samples of C_{max}	lnC_{max}
160.4371	5.077902		10.2324	2.325559
224.7913	5.415172		9.3668	2.237172
132.1218	4.883724		12.6993	2.541547
196.481	5.280566		8.5327	2.143906
174.1841	5.160113		9.4104	2.241815
153.6179	5.034468		10.7164	2.371775
165.4031	5.108385		9.7471	2.27697
222.1206	5.403221		15.699	2.753597
160.4371	5.077902		16.2239	2.786485
222.1206	5.403221		11.5974	2.450781
214.6501	5.369009		17.303	2.85088
214.6501	5.369009		10.7164	2.371775
187.9177	5.236004		21	3.044522
214.6501	5.369009		14.7725	2.692767
171.7206	5.145869		12.8171	2.55078
132.1218	4.883724		11.2346	2.418998
222.1206	5.403221		11.4669	2.439465
165.4031	5.108385		9.3668	2.237172
140.0086	4.941704		16.2239	2.786485
153.6179	5.034468		8.5327	2.143906
207.7972	5.336562		7.3284	1.991757
167.3105	5.119851		8.4611	2.135479

Bootstrap Samples of $AUC_{0-\infty}$	$lnAUC_{0-\infty}$		Bootstrap Samples of C_{max}	lnC_{max}
140.0086	4.941704		11.5974	2.450781
167.3105	5.119851		8.4611	2.135479
140.0086	4.941704		9.7471	2.27697
119.3674	4.782206		11.4642	2.439229
153.6179	5.034468		10.3462	2.336619
187.9177	5.236004		15.699	2.753597
222.1206	5.403221		8.0715	2.088339
167.3105	5.119851		11.5974	2.450781
167.3105	5.119851		7.3284	1.991757
196.481	5.280566		8.0715	2.088339
165.4031	5.108385		13.5789	2.608517
202.519	5.310834		11.5974	2.450781
173.7069	5.15737		12.7964	2.549164
154.1307	5.037801		11.4669	2.439465
184.5847	5.218108		11.5974	2.450781
132.1218	4.883724		7.3284	1.991757
154.1307	5.037801		16.2239	2.786485
183.2101	5.210633		14.7725	2.692767
171.7206	5.145869		12.6993	2.541547
222.1206	5.403221		13.5789	2.608517
167.3105	5.119851		12.8158	2.550679
207.7972	5.336562		9.5904	2.260763

Bootstrap Samples of $AUC_{0-\infty}$	$lnAUC_{0-\infty}$		Bootstrap Samples of C_{max}	lnC_{max}
153.6179	5.034468		10.2324	2.325559
183.2101	5.210633		10.7164	2.371775
207.7972	5.336562		12.8171	2.55078

Appendix 10 SAS codes

Power *AUC*

```
data pairedauc;
Mu1=5.1432; Mu2=5.1566; StDev1=0.04412;StDev2=0.04412; Corr=0.691717;
N=28;Alpha=0.05;
StDevDiff = sqrt(StDev1**2 +StDev2**2 -2*Corr*StDev1*StDev2);
NCP = (Mu2-Mu1)**2 /(StDevDiff**2/N);
CriticalValue = FINV(1-Alpha, 1,N-1, 0);
Power = SDF('f', CriticalValue,1, N-1, NCP);
proc print data=pairedauc;
run;
```

Power calculation

```
proc power;
twosamplemeans test=equiv_diff alpha=0.05
lower=&log_pt_8 upper=&log_1_pt_25 std=&std_derived
meandiff=&log_true_gmr
npergroup=.
power =0.95;
run;
```

Power curve

```
proc power;
twosamplemeans test=diff
nulldiff= 0
meandiff= 4
stddev= 3
power= 0.8 0.9 0.95
alpha = 0.05
npergroup = .;
plot y=power yopts=(ref=0.8 0.9 0.95);
run;
```

ANOVA

lnAUC

```
proc glm data=lnAUC;
  class Treatment Period Seq Animal;
  model lnAUC=Treatment Period Seq Animal;
  random Animal/ test;
run;
```

Mixed Procedure: ANOVA

lnAUC

```
proc mixed data=lnAUC;
     class Seq Period Treatment Animal;
     model lnAUC=Seq Period Treatment;
     random Animal(Seq);
     lsmeans Treatment/pdiff cl alpha=0.1;
     estimate 'T/R' Treatment 1 2 / cl alpha=0.1;
    * make 'LSMEANS' out=lsmean; *used in old SAS versions;
    * make 'estimate' out=est; *used in old SAS versions;
run;
```

ANOVA code for covariances

lnAUC

```
proc mixed data=covauc;
classes Seq Animal Period Treatment;
model lnAUC= Seq Period Treatment / ddfm=satterth;
random Treatment/type=FA0(2) sub=Animal G;
repeated/grp=Treatment sub=Animal;
estimate 'test-ref' Treatment -1 1/ CL alpha=0.10;
run;
```

TOST

lnAUC

```
ods graphics on;
proc ttest data=TOSTFL dist=lognormal tost(0.8, 1.25);
paired TestlnAUC*ReflnAUC;
run;
ods graphics off;
```