A Comparison of classical Bioequivalence analysis techniques with Simulated annealing algorithm optimised sampling times and population pharmacokinetic modelling

School of Statistics and Actuarial Science

BY

Senyo Frank Kofi Cudjoe

937561

Supervisor

R Krommenhoek

A Dissertation submitted to the Faculty of Science,

University of the Witwatersrand, Johannesburg, in

fulfilment of the requirements for the degree of Master of

Science

October, 2017

Declaration

By submitting this dissertation for the qualification of Master of Science degree in Mathematical Statistics at the University of the Witwatersrand, I declare that the entirety of the work contained therein is my own, original work and that I am the authorship owner thereof (unless to the extent explicitly) and I have not previously in its entirety or part submitted it for obtaining any qualification at another university. I further declare that all sources cited or quoted are indicated and acknowledged by means of a list of references. I further cede copyright of this dissertation to the University of the Witwatersrand.

Signed:
_____

Date:
_____

Abstract

Bioequivalence (BE) studies are conducted to demonstrate that two drug formulations produce similar bioavalabilities or therapeutic effect and safety when used. During this study, drug concentrations are obtained several times over a given period of time and a concentration vs. time graph is constructed. For a generic drug to be approved, this BE studies are conducted and the generic drug must be demonstrated to be therapeutically equivalent to the innovator drug.

This study utilised a standard 2×2 crossover design to randomly assign subjects to each of the two sequences. Statistical methods such as confidence intervals, Schuirmann's two one-sided and Wilcoxon-Mann-Whitney tests were used to assess average bioequivalence (ABE). However, there are concerns that the use of ABE alone is not appropriate for drugs with high intra-subject and inter-subject variabilities. Under such a circumstance, population bioequivalence (PBE) and individual bioequivalence (IBE) are proposed. This study employs the PBE approach but not the IBE as it is not possible to perform IBE on the available data. Results indicated that the generic drug is average bioequivalent to the innovator drug although C_{max} was outside the regulatory range set by the Food and Drug Administration (FDA).

Most biological data are modelled using nonlinear fixed effect models. Population pharmacokinetic (PK) modelling has been used in clinical pharmacology to identify the sources of PK variability in the target population. This study was conducted to determine the characteristics of the PK parameters of the orally administered antibiotic given to pigs using a population approach. A population PK model was developed using a nonlinear mixed effects model (NLMEM) with a one-compartment model using different residuals. For the NLMEM, the stochastic approximation expectation maximisation (SAEM) algorithm was implemented in MONOLIX. The models were used to estimate

the population PK parameters and diagnostic plots obtained for model evaluation. The results showed that the combined residual error model fitted the data better than the constant error model.

In addition, this study sought to find optimal sampling times which will minimise the number of blood samples required for pharmacokinetic study. The optimal sampling times were generated from a one compartment model and implemented in MATLAB. The parameters used in the optimisation were estimated from the population PK model. These sampling times were generated using the simulated annealing (SA) algorithm.

Keywords: Bioavailability, Bioequivalence, generic drugs, reference drugs, average bioequivalence, population bioequivalence, individual bioequivalence, pharmacokinetics, pharmacodynamics, therapeutic window, confidence interval, nonlinear fixed effects model, two one-sided tests, maximum likelihood, population parameters, optimal sampling design, concentration-time curve, model evaluation, Visual predictive checks (VPC), Normalised prediction distribution errors (NPDE), Stochastic approximation expectation maximisation (SAEM) algorithm.

Acknowledgements

Dedication

In memory of my grandmother

Naomi Kokui Akaba Dzakpasu

1903-2001

Contents

List of Figures

List of Tables

List of Acronyms

Acronym	Description
ABE	Average Bioequivalence
AIC	Akaike Information Criterion
ANDA	Abbreviated New Drug Application
ANOVA	Analysis of variance
AUC	Area under concentration time curve
BA	Bioavailability
BE	Bioequivalence
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CI	Confidence Intervals
CMAX	Maximum drug concentration
COD	Crossover Design
CTC	Concentration Time Curve
CV	Coefficient of variation
DF	Degrees of freedom
EBE	Empirical Bayes Estimates
ED	Estimates Distance
EM	Expectation Maximisation
EMEA	European Medicines Evaluation Agency
FDA	Food and Drug Administration
FO	First Order
GOC	Global Optimality Criterion
HT	Hotelling T^2 Statistic
IBE	Individual Bioequivalence
IDR	Individual Difference Ratio

List of Acronyms

Acronym	Description
IV	Intravenous
IWRES	Individual Weighted Residuals
LD	Likelihood Distance
LMEM	Linear Mixed Effect Model
LRT	Likelihood Ratio test
MA	Metropolis Algorithm
ML	Maximum likelihood
MM	Method of Moments
MMAP	Maximum Medical Aid price
MS	Mean square
MSE	Mean Square Error
NCA	Non compartmental analysis
NLMEM	Nonlinear mixed effects models
NLMM	Nonlinear Mixed Model
NPDE	Normalised prediction distribution errors
OTA	Office of Technology Assessment
PD	Pharmacodynamics
PDR	Population Difference Ratio
PBE	Population Bioequivalence
PI	prediction Interval
PK	Pharmacokinetic
PSSA	Pharmaceutical Society of South Africa
PWRES	Population Weighted Residuals
R	Reference
REML	Restricted Maximum Likelihood
RSE	Relative Standard Error

List of Acronyms

Acronym	Description
SA	Simulated annealing
SAEM	Stochastic approximation expectation maximisation
SAMCC	South African Medicine Control Council
SAS	Statistical Analysis System
SE	Standard Error
SS	Sum of Squares
STS	Standard two Stage
T	Test
TOST	Two One-Sided Test
USA	United Sates of America
VPC	Visual predictive checks
WHO	World Health Organization

List of Symbols

Symbol	Description
AUC_{0-t}	Area under a curve from zero to time t
$AUC_{0-\infty}$	Area under a curve from zero to infinity
Cl	Drug clearance
C_{max}	Maximum drug concentration
D	Dose of administered drug
K_a	Absorption rate constant
K_e	Elimination rate constant
T_{max}	Time of maximum concentration
$T_{1/2}$	Half-life of a drug
V	Volume of Distribution
ξ	Residual error parameter

Chapter 1

Introduction

1.1 Background of the study

Some patients live most of their lives on drugs in order to control certain conditions or cure illnesses. Some of these patients suffer from conditions such as epilepsy or asthma which require regular use of drugs. These patients take drugs to relieve pains which they feel most often. Whenever drugs are prescribed by medical doctors and given by pharmacists to subjects, a schedule of the quantity and time during which the drugs should be taken is prescribed and generally referred to as dosage regimen. Dosage regimen refers to a schedule of doses of a therapeutic agent per unit of time including the time between doses, for example, every 2 hours or time the doses are to be taken, for example, 9am and 5pm daily, or the amount of medicine, for example, number of capsules, to be given at each specific time (Rowland and Tozer, 1995). The duration of the drug therapy and the dosage regimen depends on the objective which could be a prevention of the disease, cure or reduce the severity of the disease.

Answers to therapeutic questions were obtained using a trial and error approach in the past (Rowland and Tozer, 1995). With the trial and error approach, the dose, the interval between the dose and the route of administration are selected by the pharmacist. The medical doctor then follows the effect of

the drug on the subject, takes note of the desired effects of the drug such as pain relief and any complex issues arising from it. The dosage regimen may then be adjusted until a suitable balance is obtained between the desired effects and the dose of the drug which may be harmful to the body. After a substantial experimentation on a large number of subjects, a dosage regimen which seems to be reasonable is established. In order to prevent the drugs from being toxic to the body or being ineffective when taken when using the trial and error approach, it is very important to examine what happens after a drug is administered to a subject. In order to administer a drug optimally, one needs to understand or have knowledge of the mechanisms of the drug absorption, distribution, elimination and also the kinetics of these processes, i.e. pharmacokinetics (Rowland and Tozer, 1995).

Drug administration can be divided into two phases, namely pharmacokinetic phase and pharmacodynamic phase (Rowland and Tozer, 1995). In the pharmacokinetic phase, dose, dosage form (tablet, capsule, solution, etc), frequency and route of administration are related to drug level-time relationships in the body and in the pharmacodynamic phase, concentration of drug at the site(s) of action is related to the magnitude of the effect(s) produced (Rowland and Tozer, 1995). Pharmacokinetics (PK) is the study of the time course of the absorption, distribution, metabolism and excretion (ADME) of a drug after its administration to the body (Bauer, 2008). PK deals with the movement of the drug in a body after its administration while pharmacodynamics (PD) is the study of the relationship between the concentration of a compound at its site of action, where the therapeutic targets are located and the magnitude of the pharmacological response (Fan and de Lannoy, 2014). In simple terms, PK describes what the body does to the drug while PD describes what the drug does to the body (Abel-Rahman and Kauffman, 2004).

After a drug is administered into the body, it normally undergoes four different stages before being eliminated from the body. These four stages are

known as ADME: Absorption (process where the drug is taken by the body through the mouth), distribution (the process where the drug is carried by the body through blood to action site), metabolism (the breaking of the drug substance into by-products), and elimination (dispersion of the drug product by the body).

In PK studies, after a drug is administered, the subjects are monitored for a period and a certain number of blood samples are drawn to measure the time-course of the plasma drug concentration (Choi, Caffo and Rohde, 2007). An approximation of the concentration-time curve (CTC) is constructed. PK parameters such as area under the curve (AUC), maximum concentration (C_{max}) and time to reach maximum concentration (T_{max}), are then estimated using the observed drug concentration profile for each person. Optimisation of the sampling times increases the accuracy of the PK parameters and reduces the number of blood samples which are drawn from the subjects (Asyali, 2010). This enables an optimum number of blood samples to be taken, which may give better results and prevent inconvenience on the part of the subject.

Since the basic concepts of PK are common to all drugs, information from the PK of one drug can be used in anticipating or projecting the pharmacokinetics of another drug. Knowledge of the pharmacokinetics of a drug helps the pharmacist to predict the optimal dosage regimen for individual patients and also what happens when a dosage regimen is altered (Rowland and Tozer, 1995). Whenever a new drug is discovered, an appropriate dosage form must be developed for the drug to be distributed efficiently to the body so that the body can obtain beneficial results from the treatment (Chow and Liu, 2008). Different dosage forms could be designed for a drug to serve a specific purpose. For example, liquid dosage forms may be appropriate and convenient for different categories of individuals such as children and elderly people. Generally, most drugs are administered orally which are often in solid forms. These solid drugs have to be dissolved in the blood or the tissues in the body after

administration of the drugs in order to be absorbed by the body.

1.2   Definition of Bioequivalence studies

Bioequivalence (BE) studies are conducted by a generic pharmaceutical industry as well as pharmaceutical companies that seek access to market generic drugs of brand formulations when their patent expires. The PK parameters of the CTC are used to demonstrate the safety and efficacy of the generic drug (bioequivalence) instead of performing a clinical trial which is usually more expensive and time consuming. These studies are usually conducted to reveal that two drugs are similar to each other in terms of safety (non-therapeutic side effects) and efficacy (therapeutic benefit) and therefore very important in drug development (Patterson, 2003). They are carried out to meet regulatory standards that are set by bodies such as the Food and Drug Administration (FDA), European Agency for the Evaluation of Medicinal Products (EMEA) and the South African Medicine Control Council (SAMCC) and are the main focus of this dissertation.

A bioequivalence study is performed on the test drug to assess the clinical or psychological effect. The motivation for this study is to assess the expected in vivo biological equivalence of two proprietary preparations of a drug. If two products are bioequivalent, it means that they would expect to be, for all intents and purposes, the same.

1.3   Statement of the Research problem

When the right to produce and sell a drug expires, the company that developed the original drug (a base measurement for others) may try to develop a new formulation with the same active ingredients (Rebull, Sanchez, Pla and Pla, 2008). During the development period of the new drug formulation, BE studies is used to address the issue of high costs of drugs as this study produces

drugs which cost less compared to the reference drug. This is mostly done to reduce the price of the drugs in the market and thereby making it affordable to the general public whilst serving the same purpose as the reference drug. Other companies may also try to develop a generic drug that is comparable in dosage form, quality and strength to the original drug.

A generic (test) drug is a drug which has more or less the same chemical composition and is as effective, after it has been administered to a patient, as the original (reference). To obtain approval to produce the test drug, the FDA, SAMCC and other agencies require proof of equivalence between the reference and test drug. The test drug, however, ought to be identical to the reference drug in terms of efficacy, safety and usage. For the test drug to be accepted and sold in the market, it needs to meet standards set by the FDA and other agencies.

This concept, which is termed BE, is used to describe two drugs with two different dosage forms, yet have similar bioavailability and are pharmaceutically equivalent both in terms of efficacy and safety after administration (Chow and Liu, 2008). BE attracts global attention particularly in the pharmaceutical industry but is also of major concern to government agencies, healthcare providers and individuals. BE addresses a typical question: "Can a test drug be prescribed for a subject who previously has been receiving the reference drug without any significant change in therapeutic effect"?

A BE study is performed on the test drug to assess the clinical or psychological effect. This study is done to enable drug switching to reduce the cost of the drugs and to compare the test drug's performance to that of the reference drug. The motivation for this study is based on the urgent need to address drug affordability as well as to have test drugs that have the same therapeutic effect as the reference drugs.

The concept of bioavalability (BA) refers to the rate and extent by which the drug is available at the site of action (Chow and Liu, 2008; Rebull *et al.*, 2008). BA is expressed by several parameters obtained from the curve of the concentration-time graph of each subject after a single dose administration of a drug. The main BA parameters are T_{max}, C_{max}, AUC_{0-t}, and $AUC_{0-\infty}$ (Chow and Liu, 2008).

The problem can be addressed by designing a study to compare formulations of a reference drug with a generic drug with respect to pharmacokinetic parameters such as AUC, C_{max} and T_{max}. It is also desirable to determine the optimal times blood samples are collected from subjects for the purpose of having a better and precise estimate of the PK parameters and possibly reduce the cost of the study and the inconveniences brought upon the subjects in the study. During the design phase of BE studies, blood samples are drawn over time from subjects administered with drugs and PK parameters are estimated. The pharmacokinetic parameters often estimated or measured are AUC, C_{max} and T_{max}. Based on the values of the PK parameter estimates, the drug could be accepted as the test drug for the reference drug or rejected. The main aim of this study is to determine if bioequivalence can be obtained using different classical approaches and to determine the optimal sampling times for such a study. This helps to reduce the cost of the bioequivalence study and reduce the inconveniences to the subjects on whom the study is conducted. The findings will help pharmaceutical companies to manufacture test drugs which can be used to treat the same or similar conditions the reference drugs could do but being more cost effective.

In most cases, modelling biological data involves fitting nonlinear models to the available data. In PK, a subject's reaction to a drug is studied which usually is a quick increase in the concentration after the drug is given and a gradual nonlinear decrease as it moves through the body to the action site and eventually eliminated from the body. There are a number of methods

and software which are used in modelling population PK but are not able to diagnose the model fit properly. This is addressed in this dissertation through developing different models using a one-compartment model and checking the model assumptions and goodness of fits for the models and visualising the model diagnostics. Parameter estimates were determined which were used from the optimal sampling times to determine bioequivalence.

1.4 Significance of the study

The rate and extent of BA could be affected or influenced by any small change in the contents of the formula and compaction of the drug into tablets (Patterson, 2003). The use of test drugs as substitutes for the reference drugs enables the market to be provided with inexpensive, efficacious and safe drugs without necessarily repeating the entire clinical development of the drug, which drives the need to conduct BE studies.

The cost of healthcare has been increasing over the past decades and the price of drugs was identified as the main cause (Chow and Liu, 2008; Midha and McKay, 2009; Kamerow, 2011). According to Gray (2009), a commission of inquiry was established in South Africa in 1961, 1978 and 1985 to investigate the escalating costs of healthcare which included medicines. The recommendations of the three commissions of inquiry, namely the Snyman, Steenkamp and Browne commissions, stated that patent legislation was the main cause of high prices of drugs in the South African pharmaceutical market and identified test drugs as a way to save the cost of healthcare (Gray, 2009).

The reference and test drugs are both available in the market; however, the reference drugs are more expensive than the test drugs. The availability of the test drug in the market helps to create competition between the reference and the test drug which eventually results in a significant drop in the price of the drugs resulting in making the drug affordable. However, for the test drug

to be accepted and sold in the market, it needs to meet regulatory standards which are set by regulatory bodies such as FDA, MCC and others.

According to Henry and Lexchin (2002), lack of access to the most important drugs is not only a problem that faces less developed countries but also people in the developed countries. Some of the people, especially elderly people and people without medical insurance, are not able to afford the drugs they need due to the high prices of these drugs. The introduction of the test drug form of Omeprazole, as stated by Henry and Lexchin (2002), in the Australian market saw a significant drop of 43% in the price of the reference drug Losec within a period of 2 years and also a 97% drop in the price of the test drug form of the combination antiretroviral drugs after being marketed by a test drug producer. These significant price drops in the prices of the two drugs are illustrated in Table 1.1.

Table 1.1: Effect of test drugs on the prices of drugs.

Omeprazole(Australian)				Antiretroviral Combination Therapy			
Generic product		Branded product		Generic product		Branded product	
Date	Price (US$)	Date	Price (US$)	Date	Price (US$)	Date	Price (US$)
November 1998	N/A	November 1998	43.28	July 2000	230.59	September 2000	869.92
November 1999	30.86	November 1999	31.27	September 2000	66.67	October 2000	77.58
November 2000	30.01	November 2000	30.63	February 2001	29.17	March 2001	59.33
November 2001	24.10	November 2001	24.51	August 2001	24.58	August 2001	59.33

The motivation for this dissertation is based on the urgent need to address the issue of drug affordability as well as to have test drugs with similar therapeutic

effects as the reference drugs. The findings will help inform pharmaceutical companies to manufacture test drugs which can be used to treat similar conditions to those treated by the reference drug but at an affordable price.

According to Karim, Pillai, Ziqubu-Page, Cassimjee and Morar (1996), the escalation in the health care cost in South Africa is partly due to the high costs of drugs. They further stated that test drugs have been identified as a mechanism for healthcare cost reduction. This has led the Pharmaceutical Society of South Africa (PSSA) to introduce the Maximum Medical Aid Price (MMAP) as a payment option in the private sector to encourage the use of test drugs. The study reviewed 1570 prescriptions which had 4086 items and found that 45.7% of the total prescriptions had at least an item with a test drug equivalent. However, only 0.3% of the prescribers were against test drug substitutions. The reference drugs' prices were 9.9% greater than their test drugs' prices (Karim *et al.* 1996) as indicated in Table 1.2.

Table 1.2: Prices of prescription (in Rands) received by 10 Pharmacists on 4 selected days.

	Price of all 1570 Prescriptions	Mean price
Reference drug	189042.56	120.49
Test drug	172038.68	109.65

1.5   Aims

The aims of the study are to:

- Determine the optimal times blood samples could be taken from subjects after administration of test drug formulations in order to estimate PK parameters such as AUC, C_{max} and T_{max};

- Establish bioequivalence between the reference (R) and test (T) drug formulations for both the original measurement and the simulated results; and

- Do population Pharmacokinetic modelling.

1.6 Objectives

The key objectives of the study are to:

- Test for carryover effect, period effect and direct drug effect;

- Perform Analysis of variance (ANOVA) and use the results from the ANOVA in the confidence interval to determine if the two formulations are bioequivalent;

- Perform a Schuirmann's two one-sided t-test (TOST) to assess average bioequivalence (ABE) between the two formulations for AUC and C_{max};

- Carry out a non-parametric test i.e Wilcoxon-Mann-Whitney test on the T_{max};

- Perform population bioequivalence;

- Determine the power of Schuirmann's TOST for the hypothesis of formulation effects, determine the size effect of the TOST;

- Use the nonlinear mixed effect model to develop a population pharmacokinetic model using a one-compartment model and estimate population pharmacokinetic parameters;

- Determine optimal sampling times using simulated annealing algorithm;

- Perform BE studies using the optimised sampling times and the concentrations; and

- Compare the results of the classical methods obtained from the original data to the optimised data.

1.7   Dissertation Organisation

This dissertation consists of five chapters. The next chapter (Literature Review) provides details of the history of BE studies and some of the methods used in the assessment of bioequivalence studies. A further background of optimisation techniques which are used to obtain optimal sampling times in BE studies are reviewed in this chapter. The research methodology as well as the optimisation method for estimating sampling times using simulated annealing algorithm are explained in Chapter 3. The results of the methods described in Chapter 3 are presented and discussed in Chapter 4. Chapter 5 discusses the conclusion and provides recommendations.

1.8   Limitations of the study

In this dissertation, the 2×2 crossover design which is widely used was adapted even though it is not optimal. Higher crossover designs usually give better results compared to the 2×2 design as they test for within and between subject variability. The higher crossover design was not available for the data used in this dissertation. This therefore led to the inability of performing IBE which requires the higher order design.

The cost of the reference and test drugs were also not known so comparison to ascertain if the test drug is cheaper than the reference drug was not carried out.

Other characteristics of the subjects such as weight and gender were not available for the population modelling and hence these covariates were not included in the modelling.

Chapter 2

Literature review

2.1 Introduction

The theory of BA and BE have attracted so much attention from academia, health authorities and pharmeceutical industries over the past 30 years due to its importance and implementation to test drugs (Hauschke, Steinijans and Pigeot, 2007). This research is necessary due to the high cost of healthcare particularly the high cost of drugs. This has prompted regulatory bodies to start formulating requirements that are necessary for approving test drugs as substitutes for the reference drugs. In the early 1970s, BE was generally tested using the 75/75 rule (Hauschke *et al.,* 2007). This rule used the ratio of AUC of the test drug formulation to the AUC of the reference drug formulation in the BE study and states that it must be between 75 and 125 per cent in order to declare that the two drugs are bioequivalent (Buehler and Director, 2010). However, this approach of assessing the bioequivalence between two formulations received a lot of criticism. It was therefore used together with the power approach in the early 1980s. The power approach is a method of testing interval hypothesis of no difference using a standard two-sided t test at the 0.05 level of significance (Schuirmann, 1987). If the null hypothesis of no difference under the power approach is rejected, then μ_T and μ_R are not considered to be equivalent and if the hypothesis of no difference is not rejected, the power of the test is being questioned. The power should be at least 0.80 before failure

to reject the hypothesis of no difference may be taken as evidence that μ_T and μ_R are equivalent.

According to Midha and McKay (2009), the FDA started to research the BA of new drugs which led to the formation of a BE study panel by the Office of Technology Assessment (OTA). This office was mandated to study the therapeutic and chemical equivalence of two drug formulations. The United States of America congress then enacted the Drug Price Competition and Patent Term Restoration Act of 1984 authorising the FDA to accept test drugs for use when there is evidence of BA and BE studies (Midha and McKay, 2009). The activities of the FDA are the examination, acceptance or rejection of the test drugs application. A core objective of the act was to ensure that even though less expensive test drugs are produced, the quality was not compromised so as to ensure the safety and efficacy whenever it is taken by patients. The FDA frequently publishes guidelines on how BA and BE studies can be conducted on test drugs. These guidelines provide the different statistical methods that are approved by the FDA and their acceptance criteria.

BE studies are important in assessing the efficacy, safety and clinical effect of test drugs. They enable effective switchability between the reference and test drugs thereby making the test drug affordable to most people. The ability to collect blood samples from subjects at optimal times save cost of BE studies and reduce the inconvenience caused to the subjects. The next section is a review of literature on BE studies.

2.2 History of Bioequivalence

Clinical trials that are used to register drugs into the market place use small samples of the total population of people who will eventually take the drugs. Due to this, not all the patients will benefit from it and experience the same level of safety while taking the drug. However, the risk should be minimised

or held at an acceptable level with regards to the policies. When the patent period of a drug expires, new formulations (test) could be marketed which generate profit for the companies that produce them and offer benefits as well to the general public (lower costs) (Strom, 1987). However, there were reports of failures of some of the test drugs in the United States which received public attention leading to establishing standards and requirements for test drugs to be accepted into the market especially drugs which have a very narrow therapeutic index (drugs which result in large change response with a small change in dose) (Rheinstein, 1990; Calvert, 1996). There are other examples of bioinequivalence for Carbamazephine as stated in Welty *et al.* (1992).

The FDA was then given the authority by the Drug Price Competition and Patent Term Restoration Act of 1984 to generate a procedure of approving test drugs which needed to be introduced into the marketplace. The FDA in 1985 approved one hundred and twenty two test drugs whose patent periods expired for marketing (Strom, 1987). According to Patterson (2003), crossover designs were very important and of general interest in clinical trials in the 1960s through to the 1980s. In a crossover design there is a random allocation of subjects to different sequences where each sequence receives different treatments (Jones and Kenward, 2014). A further discussion on crossover design is conducted in Chapter 3.

Jones and Kenward (2014) indicated that the history of BE studies started in the 1960's. Also, bioequivalence studies, according to Midha and McKay (2009), have been ongoing for the past decades and are accepted as the measure of approval for test drugs to be marketed and sold at reduced costs. BE studies are conducted to demonstrate that drug products are equal in rate and their extent of absorption (Patterson, 2001). Patterson further stated that bioequivalence studies dated back to the 1970's through to the 1990's with average bioequivalence being the accepted standard for the approval of test drugs. Chow and Liu (2008) also indicated that bioequivalence studies, which

14

dated back to 1970's, could be divided into four stages.

The first stage, which dates back from the early 1970's to 1984, during which the Drug price Competition and Patent Restoration Act was passed, gave the authorization to the FDA to accept test drugs when there were BA and BE studies available. The second phase began from 1984 to 1992 and provided companies with guidance on how the data should be analysed and presented. The third phase started in 1992 and addressed population and individual bioequivalence concepts as well as statistical methods used for each. The fourth phase started at the beginning of the twenty-first century. It was based on research conducted in the past three decades in the twentieth century prompting the FDA to issue and implement guidelines for the statistical methods for bioequivalence studies.

Test drugs in South Africa are controlled and regulated by the Medicines Control Council which was formed in the 1970s (MCC, 2003). It was constituted under the Medicines and Related Substance Act 101 of 1965 to monitor and regulate the manufacturing of test drugs. The MCC is mandated to ensure that drugs that are marketed and used in South Africa are safe for public consumption.

Kong and Gonin (2000) use an optimisation technique for calculating the optimal time design for one compartment model. This was to select the sampling times optimally in order to improve the accuracy of the AUC estimates in a BE test. In the article they consider a compartment open model while assuming that the drug follows a first order absorption. The compartmental model in Equation 2.1 has the concentration-time curve given as

$$C_t = \frac{DK_a}{V(K_a - K_e)}(e^{-K_e t} - e^{-K_a t}), \qquad (2.1)$$

where D, V, K_a and K_e are positive constants, $0 \leq t \leq \infty$ where t is the time the blood samples are drawn, D denotes the dose of the administered

drug, the volume of drug distribution is denoted by V, and the absorption and elimination rate constants of the drug are denoted as K_a and K_e, respectively. The dose administered as well as the volume of distribution were considered to be constants for healthy subjects, whereas the elimination rate constant depended on the drug formulation and subjects, but was assumed to be constant due to its low variability.

Kong and Gonin (2000) conducted simulation studies using different values of K_a and T_f (final sampling time) and found that the proposed optimisation technique improves the regular design. From the empirical results, the unexplained AUC mean percentage for the regular design is -1.1 with a standard deviation of 1.5 while the unexplained AUC mean percentage for the optimal design is 0.001 with a standard deviation of 0.9. The study suggested six to eight sampling points to be the best number of sampling times in BE studies. The proposed optimal design method which was used in the study induces little bias in the estimation of the total AUC and the calculated AUC is more reliable.

Choi *et al.* (2007) in their article investigated the appropriate time to choose samples in order to evaluate the AUC. They explored the advantages of several objective functions and suggested an objective function which addressed the inadequacies of the existing objective functions. A simulated annealing algorithm (SA) was used in this minimisation process. A new objective function based on squared bias was used to estimate AUC as well as a Monte Carlo integration approach which was used to simplify the calculations. They discussed two methods for defining the objective functions. The first uses the mean squared error (MSE) for estimating AUC while the second uses MSE for estimating the concentration time curve using linear interpolation to define

the objective function. They proposed an objective function

$$Obj_G(t; \beta) = \sum_{j=1}^{m+1} \left\{ \left[E(\hat{A}_{t_{j-1}}^{g(j-1)(1)}) - A_{t_{j-1}}^{g(j-1)(1)} \right]^2 + \sum_{k=2}^{n_j} \left[E(\hat{A}_{g(j-1)(k-1)}^{g(j-1)(k)}) - A_{g(j-1)(k-1)}^{g(j-1)(k)} \right]^2 + \left[E(\hat{A}_{g(j-1)(n_j)}^{t_j}) - A_{g(j-1)(n_j)}^{t_j} \right]^2 \right\}. \quad (2.2)$$

which divides the number of divisions between the time intervals t_{j-1} and t_j equally based on the interval length. This makes it possible to estimate the errors made in approximating the AUC and the true AUC for every division after which the results are to be summed up.

Gibaldi and Perrier (1982) used a one-compartmental pharmacokinetic model which has a first-order elimination kinetic and first-order absorption which is similar to the one in Equation 2.1.

Asyali (2010) used the cubic spline approximation approach to the concentration-time curve and initiated a global optimality criterion whose main focus was to determine how close all the PK parameter estimators were to their true values simultaneously. He investigated a one compartmental model with a first order absorption rate which is the same as the one in Equation 2.1. Asyali (2010) proposed the global optimality criterion (GOC) which uses the euclidean norm as a measure of closeness of the estimated values to the actual values to indicate the accuracy of AUC_{0-24} (area under the concentration time curve from zero to 24 hours), C_{max} (peak drug concentration) and T_{max} (time required to reach peak drug concentration). The GOC is given as $GOC(t_1, t_2, ...t_N) = \| \hat{\theta} - \theta \|^2 = \left(A\hat{U}C_{0-24} - AUC_{0-24} \right)^2 + (\hat{C}_{max} - C_{max})^2 + (\hat{T}_{max} - T_{max})^2$. The optimal sampling times are determined by minimising the GOC with respect to $t_1, t_2, ...t_N$, where N is the last sampling time. In the study, he considered several ways to enhance the spline approximation in order to obtain the optimal sampling times by minimising the GOC using another optimisation technique known as sequential quadratic programming (SQP). The SQP is a method for numerical solution of constrained nonlinear optimisation problems. It is a conceptual

method from which numerous specific algorithms have evolved and is backed by a theoretical and a computational foundation. The SQP method finds an approximate solution of a sequence of quadratic programming subproblems in which a quadratic model of the objective function is minimized subject to the linearised constraints. It is established that this approach gives more precise estimates of PK parameters with a small number of samples.

Jones *et al.* (1999) described a method which is used to construct D-optimal designs for a nonlinear mixed effect model (NLMEM) in crossover PK studies based on a model which uses generalised least squares. The variables, such as the parameter values as well as the design points, are specified in the model while samples are simulated under the simulation approach. The performance of the design is assessed using the estimated model parameters and their corresponding variances. A nonlinear mixed model (NLMMs) in Equation 2.3 with random and fixed effects was used to analyse the data,

$$y_{ijq} = f(t_q, \theta_{ij}) + \varepsilon_{ijq}, \tag{2.3}$$

with y_{ijq} being the concentration for the ith subject in period j at time q, $f(t_q, \theta_{ij})$ is a nonlinear function representing the mean pharmacokinetic response, the measurement time represented as t_q, and θ_{ij} represents a vector of pharmacokinetic parameters for the ith subject in the jth period. $\varepsilon_{ijq} \sim N(0, \sigma_e^2)$ is a random measurement error. The pharmacokinetic parameters were estimated using Equation 2.4:

$$\theta_{ij} = X_{ij}\boldsymbol{\beta} + \boldsymbol{\mu_i}, \tag{2.4}$$

where $X_{ij} = x_{ij} \otimes I_m$, I_m is an $m \times m$ identity matrix, x_{ij} represents a $1 \times l$ covariate vector which constitutes the crossover design, $\otimes$ represents the Kronecker product, $\boldsymbol{\beta}$ represents an $l_m \times 1$ vector of coefficients for the period, mean, and treatment effect and $\mu_i \sim N(0, \sigma^2)$ is a subject effect vector.

Jones *et al.* (1999) used the D-optimality criterion which makes the determi-

nant of the information matrix to be as large as possible. The values for the mean, period, treatment effect, and variability of the data are determined using Equation 2.3 to fit the within-subject model to each individual subject and Equation 2.4 to estimate parameters. The optimal design criterion changed significantly when period effects were included or excluded.

D'Argenio (1981) used a procedure which adopts a sequential estimation approach for the optimal selection of the sampling times to estimate the parameters of the model. This technique used parameter estimates from previous subjects to estimate the optimal times for the next subject in the group. Monte Carlo simulations were used to compare the estimated parameters using the optimal sequential estimation method for the parameter estimates versus using the conventional method of choosing the sampling times. He assumed that the kinetics of the drug can be represented by a first order ordinary differential equation given in Equation 2.5:

$$\dot{x} = f(x, \alpha^*, r, t), \ \ x(0) = g(\alpha^*), \tag{2.5}$$

where x is a vector representing compartmental concentration, r represents the piecewise constant input vector, the independent variable is represented as t, α^* denotes the true but unknown constant parameter vector, and $x(0)$ represents the initial condition vector which could be a function of the unknown parameter as well. From the results obtained by D'Argenio(1981), the mean and standard deviation of the optimal sampling experiment were smaller than that of the conventional sampling type (sampling points chosen randomly) from empirical results obtained. The optimal method produced smaller parameter estimate standard deviations as compared to the conventional approach. The average percentage bias was 18% for every parameter when using the optimal sampling approach while it was 56.5% for both K_{el} (elimination rate constant) and K_{cp} (central to peripheral rate constant) (D'Argenio, 1981). The conventional method gave a larger average percentage bias compared to the optimal method.

2.3 Bioavailability and Bioequivalence

2.3.1 Introduction

Bioavailability and bioequivalence are important to both the producer of the reference drug and the test drug for marketing the drug. FDA approval of any test drug requires proof of bioequivalence, usually ABE between the reference and test drug.

2.3.2 Bioavailability(BA)

According to Chow and Liu (2008), BA is the amount of a substance that becomes available to an organism's body for bioactivity when introduced through ingestion, inhalation, injection or skin contact. Rate of bioavailability depends on factors such as type of the substance, whether fat soluble or water soluble and the composition of the diet. Whenever a drug is administered either orally or through any other route, an adequate quantity of the administered drug is absorbed over a period of time before its effect can be felt. The BA for oral drugs that are delivered into the body system can be considered a measure of both the rate and total amount of drugs reaching the general circulation from an administered dosage form (Welling, 1984). This assesses the process whereby the drug is delivered from the dosage form and moves to the part of the body where the drug effects are expected and therefore the drug's absorption, distribution, metabolism and the elimination processes can be determined (Kimura and Higaki, 2002).

The major parameters measured in BA are (Terry *et al.,* 1982):

- C_{max} - the peak plasma drug concentration used to determine the rate of drug BA;

- AUC - the area under the concentration time curve which is a measure of the extent of drug BA; and

- T_{max} - the time required to reach maximum drug concentration after the administration of the drug.

The extent and rate of oral absorption of a drug is vital in pharmacokinetics as it is responsible for the physiological action of the drug and thus referred to as drug bioavailability (Caccia and Garattini, 1990). A certain fraction of the drug is metabolised while passing through the liver or other organs. Only a portion of the consumed drug reaches the site of action and this can result in ineffectiveness of the drug in treating the illness. A comparative BA study compares the BA of different drug formulations of drugs which are similar (Chow and Liu, 2008).

2.3.3 Bioequivalence (BE)

Two drugs are considered as pharmaceutically equivalent whenever they contain interchangeable quantities of similar active ingredients but not necessarily of the same amount or dosage (Chow and Liu, 2008). Comparing the therapeutic performance of two drug formulations is important to assess the chances of using a similar drug product (test) as an alternative to another (reference). These two drugs could either be produced by the same manufacturer or by different manufacturers. The BE can be measured in at least three ways: chemically, biologically, or therapeutically (Asyali, 2010).

The BE assessment is an issue which is of great concern to the biopharmaceutical industry (Metzler, 1974). A test drug may be used as a substitute for the reference drug if it complies with regulatory requirements and meets certain criteria in a particular country. The use of pharmacokinetic concepts and parameters have made BA and BE studies now acceptable for expensive, complicated and lengthy clinical trials.

There are three reasons for conducting BE studies. These are:

- when the dosage form that is proposed for the market is totally different from the one used during the clinical trials;

- when there are important changes during the process of manufacturing

a marketed formulation; or

- where a new test formulation is being tested against the reference formulation (Chow and Liu, 2008).

Ghosh and Rosner (2007) explained that whenever the rate and degree of absorption of the reference formulation and its test formulation show similar drug concentration-time profiles, they exhibit similar therapeutic effects. Test drugs need to show BE with the reference drugs, in vivo, in order to be accepted in the market as therapeutically effective. There are three kinds of bioequivalence methods: average bioequivalence (ABE), population bioequivalence (PBE) and individual bioequivalence (IBE) (Chow and Liu, 2008).

2.4 Decision Rules and Regulatory Aspects

2.4.1 Average Bioquivalence

The ABE is a recommended method by regulatory bodies including the FDA for comparing BA measures in BE studies (FDA, 2001). It mainly compares the population averages of a BE measure rather than the variances of the measure for the test (T) and reference (R) drugs (Chow and Liu, 2008). In order to show that an R drug and the T drug are average bioequivalent, it becomes necessary to prove that the C_{max} and AUC for the T drug is not significantly different from the C_{max} and AUC of the R drug.

However, ABE does not consider the variability of the T drug and R drug. The advantage of ABE is that it is easier to interpret to patients, pharmacists and physicians (Ghosh and Rosner, 2007). In a BE trial, a T drug is compared to an R drug in subjects who are healthy volunteers and have to meet certain criteria such as being older than 18 years and have a body mass index (BMI) in the range of 18.5 to 30 kg/m^2. The most frequently used statistical design for ABE comparison is the two sequence, two period crossover design (Chow and Liu, 2008). In a standard 2×2 crossover design, each subject in the s-

tudy is randomly allocated to either sequence TR (test drug in the first period followed by reference drug in the second period) or RT (reference drug in the first period followed by test drug in the second period) separated by a sufficient time period between each dosage for the drug which was administered in the first period to be totally removed from the body (Jones and Kenward, 2014). Normally at least five elimination half-lives are necessary to achieve this. A half-life is the period required for the drug concentration in the body to be reduced by one-half. Chow and Liu (2008) defined the washout period as a time interval between treatments during which the effect of one treatment is not carried over to the other. This, however, is dependent on the type of drug and the half-life, which is usually the time in which half of the drug is removed from the body. A drug formulation with a k-order carryover effect is one in which the effects of a drug administered in the first period persists up to the kth treatment period (Chow and Liu, 2008). A first-order carryover effect lasts only one treatment period. In this research, a first-order carryover effect is considered.

The ABE is determined by using the confidence interval approach or a two one-sided test procedure (Schuirmann, 1987). The ABE considers equivalence between the population means of the PK parameters for the R and T drugs. The 75/75, 80/20, 20% and 80/125 rules were proposed by the FDA for testing bioequivalence and are explained in the next sections (Purich, 1980).

2.4.1.1 75/75 Rule

Bioequivalence is declared for two drugs if at least 75% of the relative BA of the T drug to the R drug formulation of each subject falls within (75%, 125%) limits (Chow and Liu, 2008). The merits of the 75/75 rule, according to Chow and Liu (2008), are its ability to compare the relative BA within a specific subject, removal of heterogeneity of variability within subjects and applying it is easy. This rule was heavily criticised by many researchers including Haynes (1981) who revealed that the rule was responsive to certain drugs with huge

intra-subject variabilities although the mean AUC for both the T and R drug formulations may be exactly equal.

This rule may dismiss about 56.3% of the T drug as inadequate whenever the variability of the inter-subject is substantial, according to Metzler and Huang (1983). Cabana (1983) also indicated that the rule was not valid especially whenever the inter-subject coefficient of variation (CV) of the subjects involved is 60% and the intra-subject CV ranges from 20% to 30%. According to Cabana (1983), this rule works best whenever the inter-subject CV is below 40% for both the T and R drugs while the intra-subject CV is at a maximum of 30%. Also, the 75/75 rule is no more a requirement for BE assessment because it is not statistically meticulous.

2.4.1.2 80/20 Rule

A condition underlying this rule indicates that the study should be large to be able to deliver a minimum of 80% of detecting correctly any difference of 20% in the average BA (Chow and Liu, 2008). This rule relies on testing an assumption of equality for a single variable instead of testing for equivalence. When there is no statistically significant difference in the means of the R and T drugs and the power for any recognition of 20% difference of the R drug average is 80%, one can conclude that the two drugs are bioequivalent (Chow and Liu, 2008).

2.4.1.3 20% Rule

For bioequivalence to be concluded using this rule, the average BA of the T drug formulation ought to be within ±20% of the R drug. A T drug could exhibit about 20% of variability in average BA with an R drug using the ± 20% rule (Chow and Liu, 2008). However, according to Levy (1986), the ± 20% rule does not take into consideration the effect of safety and efficacy of the drug. Interchangeability of the drug formulations can be problematic when using this rule. Interchanging the T drug with the R drug can lead to a more

than 20% difference from one subject to another (Chow and Liu, 2008).

2.4.1.4   80/125 Rule

For any drug formulations to be declared as bioequivalent using this rule, the average BA of the T formulation should be within the FDA's criterion of (80%, 125%) of the R drug, mostly at 90% confidence level (Chow and Liu, 2008). It is advised that the logarithmic transformation of the AUC and C_{max} parameters should be used. However, EMEA and the World Health Organization (WHO) accept a wider confidence interval of (75%, 133%) for the C_{max} under certain conditions. The SAMCC also accepts the same wider confidence interval for the C_{max} just as the EMEA and WHO do.

2.4.2   Population and Individual Bioequivalence

The ABE for drugs with a small therapeutic index (thus drugs for which a small change in the dosage will be able to cause large changes in response to treatment) becomes very complicated (Benet and Goyan, 1995). Drugs such as Warfarin normally show very small within-subject variability (coefficients of variation less than 10%) (Patterson, 2003). However, the sample size requirement for high variabilty drugs (coefficient of variation greater than 30%) is high (greater than or equal to 30) in order to demonstrate 90% power of average BE in a two period cross-over design (Phillips, 1990). A widened equivalence limit, 0.7 to 1.43, as allowed by the EMEA, for C_{max} has been suggested by Midha, Rawson and Hubbard (1997) to allow such drugs access to the market easily. The power to demonstrate average bioequivalence reduces whenever the variability of the drug increases and it is therefore advisable to have a very large sample size so that it is sufficient to demonstrate bioequivalence.

The ABE only considers the population means of the R and T drug formulations where variation between subjects is not considered. It is therefore not able to address the concept of switchability of the drugs and therefore did not meet sufficient requirements for public consumption (Patterson, 2003). The

ABE assesses the difference in the means of the formulations and does not account for the variance of narrow therapeutic drugs and can not account for subject-by-formulation relations. Hauck and Anderson (1992) and Wellek (1993) argued that it is not enough to only show ABE in certain situations as it does not indicate relevance to the individual subject and therefore proposed the PBE and the IBE. The PBE addresses the question, "can I safely and effectively start on the R or T drug?". The IBE was introduced to answer the question "can I safely and effectively switch between the R and T drug?". For a T drug to gain approval for access into the market, PBE is required while IBE is required for switchability.
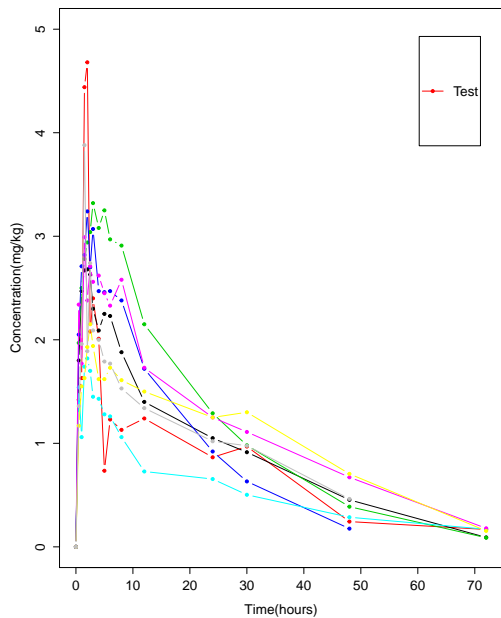
Figure 2.1: Concentration profile for T formulation in sequence 1.
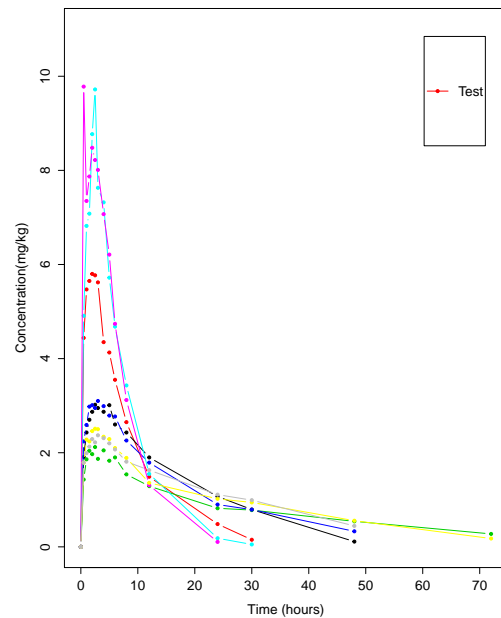
Figure 2.2: Concentration profile for T formulation in sequence 2.
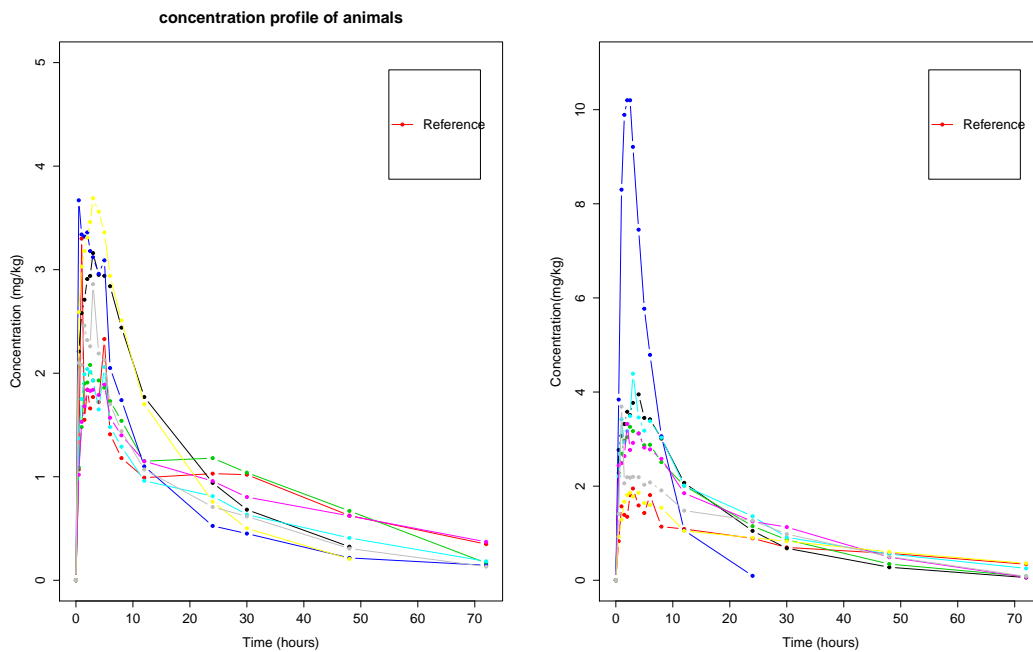
Figure 2.3: Concentration profile for R formulation in sequence 1.

Figure 2.4: Concentration profile for R formulation in sequence 2.

The T and R formulations in Figure 2.1 and Figure 2.3 show interchangeable rates of absorption and variability while the T formulation in Figure 2.2 exhibits more variability than the R formulation in Figure 2.4. According to Hauschke *et al.* (2007), the two drugs in Figures 2.2 and 2.4 can be bioequivalent though with high variability and thus can lead to different effects. ABE has not been able to address this effect leading to a new concept known as PBE which was introduced by Hauschke and Anderson (1992).

The PBE addresses the between-subject variability of the two formulations. One other concept that ABE and PBE do not address is the safety and efficacy of subjects when they switch from one formulation to another. This has to do with the subject-by-formulation interactions (Hauschke *et al.,* 2007). This concept is known as IBE and considers the subject-by-formulation interactions. The PBE and IBE both consider the means and the variability of the bioavailabilities of the subjects.

27

One of the important issues which affect BE studies is whether patients who are being treated with a T drug will experience similar results in terms of safety and efficacy as patients who are using the R drug. In practise, each subject responds differently to the same drug. Hence it becomes imperative to compare the variability of bioavailability (Chow and Liu, 2008). The exchangeability between the two drugs is questionable whenever the variability within the T drug is greater than the R drug even though the two drugs may be equivalent in average bioavailability. This is a cause of concern for the T drug. The PBE takes into account inter-subject variability (inter-subject variance) and hence solves the issue of interchangeability for subjects who need to start either treatments. If PBE is satisfied, a patient who has not received any of the formulations can be safely prescribed either the reference or test drug formulations.

After a drug is marketed, the safety and efficacy might be discussed when a drug substitution is made. This situation is related to interchangeability or switchability. IBE considers intra-subject and subject-by-formulation variances which address changes in treatment whenever an R drug is substituted by its T drug. IBE also allows for a precise evaluation of BE for drugs which have very high PK variability and with narrow or large therapeutic range (Endrenyi and Midha, 1998).

The IBE considers the possibility of switching a patient who is using the R drug to the T drug or from the T drug to the R drug (Byron and Kenward, 2003). It considers the within-subject variability of the subjects as well as subject-by-formulation relations. Chow and Liu (2008) defined an assessment criteria for PBE and IBE to be probability based and moment based as described in the next section.

2.4.2.1 Moment Based Criteria

This approach uses the loss of the expected squared error of the within subject differences of BA for subjects who receive the T and R drug formulations as well as the variability within each subject which is expressed as the loss of the expected squared error of subjects who receive the R formulation at different times (Chow and Liu, 2008). The moment based criterion consists of the ratio between intra-subject variability of the R and the within subject variability of the T drug, variability of the interaction between subject and drug formulation and the distinction in average BA (Chow and Liu, 2008). The difference ratio (DR) (Equation 2.6) was suggested by the FDA as a statistical approach for establishing bioequivalence.

$$DR = \frac{Difference \ between \ T \ and \ R \ formulations}{Difference \ between \ two \ formulations}. \qquad (2.6)$$

2.4.2.2 Probability Based Criteria

This criterion assumes that the probability of the variability due to the within subjects of BA in subjects that receive the T formulation and the R formulation on different occasions is within a pre-determined limit (Chow and Liu, 2008). However, this approach will not be considered in this research as it is not recommended by the FDA (Hauschke $et \ al.$, 2007).

2.4.3 Population Bioequivalence Using Moment-Based Criteria

The creation of the benchmark for BE requires the comparison of BA endpoints (AUC, C_{max}); differences between the R and the T as well as R with itself (Chow and Liu, 2008). Let Y_T, Y_R and Y_R' represent bioavailabilities of the T drug being administered once and the R drug administered twice. Also, $Y_T - Y_R$ and $Y_T - Y_R'$ refer to the between-subject differences when doing an assessment for PBE with Y_T, Y_R and $Y_R^{'}$ being independent (Hauschke $et \ al.$, 2007). Taking the expected squared differences and comparing the discrepancy

gives

$$E(Y_T - Y_R)^2 = E(Y_T^2) - 2E(Y_T Y_R) + E(Y_R^2)$$
$$= (E(Y_T^2) - \mu_T^2) + (E(Y_R^2) - \mu_R^2) - 2E(Y_T Y_R) + \mu_T^2 + \mu_R^2$$
$$= \sigma_{BT}^2 + \sigma_{WT}^2 + \sigma_{BR}^2 + \sigma_{WR}^2 - 2\mu_T \mu_R + \mu_T^2 + \mu_R^2$$
$$= (\mu_T - \mu_R)^2 + \sigma_T^2 + \sigma_R^2,$$

where $\mu_T - \mu_R$ represents the difference between the means of the T and the R drugs, $\sigma_{BR}^2 + \sigma_{WR}^2 = \sigma_R^2$ and $\sigma_{BT}^2 + \sigma_{WT}^2 = \sigma_T^2$, $E(Y_R - Y_R') = 2\sigma_R^2$, σ_R^2 and σ_T^2 represent the total variability for the R formulation and T formulation, respectively.

For a T drug and R drug to be declared as PBE, the equation below must be satisfied:

$$\Theta_{pop} = \frac{(\mu_T - \mu_R)^2 + \sigma_T^2 - \sigma_R^2}{max(\sigma_0^2, \sigma_R^2)} < \theta_{pop}, \tag{2.7}$$

where θ_{pop} is the predetermined value used to determine PBE for the moment-based criterion. The maximum variability difference that can exist between the T and R formulation, i.e., $\sigma_T^2 - \sigma_R^2$, is fixed at 0.02 by the FDA (1997) and σ_0^2 is fixed at 0.04.

The population difference ratio (PDR) is used to provide a motivation for the choice of 0.04 as indicated below,

$$PDR = \sqrt{\frac{E(Y_T - Y_R)^2}{E(Y_R - Y_R')^2}} = \sqrt{\frac{(\mu_T - \mu_R)^2 + \sigma_T^2 + \sigma_R^2}{2\sigma_R^2}}. \tag{2.8}$$

$$\theta_{pop} = \frac{average\ bioequivalence\ limit + variance\ allowance}{scaled\ variance}$$
$$= \frac{[\ln(1.25)]^2 + 0.02}{0.04}$$
$$= 1.74483$$

PBE is therefore concluded if $\Theta_{pop} < \theta_{pop}$. However, if the assumption of differences in the variance of the T and R formulations is not met, this criterion

of assessment of the PBE reduces to ABE. That is, if $\sigma_T^2 = \sigma_R^2$ ($\frac{\sigma_T^2}{\sigma_R^2} = 1$), then θ_{pop} is

$$\theta_{pop} = \frac{[\ln(1.25)]^2}{0.04}$$
$$= 1.24483,$$

which gives a linearized criterion of

$$(\mu_T - \mu_R)^2 + \sigma_T^2 - \sigma_R^2 - max(\sigma_o^2, \sigma_R^2)\theta_{pop} < 0$$
$$(\mu_T - \mu_R)^2 - max(\sigma_o^2, \sigma_R^2)\theta_{pop} < 0$$
$$(\mu_T - \mu_R)^2 - 0.04\frac{[\ln(1.25)]^2}{0.04} < 0$$
$$(\mu_T - \mu_R)^2 < [\ln(1.25)]^2$$
$$[-\ln(1.25)] < (\mu_T - \mu_R) < [\ln(1.25)]$$
$$0.8 < \frac{exp(\mu_T)}{exp(\mu_R)} < 1.25.$$

For any statistical assessment of PBE, the estimators of Θ_{pop} are derived using ANOVA or the restricted maximum likelihood (REML) method using mixed effects models. To solve the population bioequivalence problem, an appropriate statistical approach has to be derived:

$$H_0 : \Theta_{pop} \geq \theta_{pop}$$
$$H_1 : \Theta_{pop} < \theta_{pop}. \tag{2.9}$$

In order to conclude PBE or otherwise, a statistical approach is used where PBE can be declared when the null hypothesis shown in Equation 2.9 is rejected at the 5% level of significance. However, the FDA (1997) recommends the two-sided 90% confidence interval method or the one-sided upper 95% confidence interval method for Θ_{pop}.

The PBE looks at whether either the R or T drug can be prescribed for a subject who is not on either of the two drugs (Jones *et al.*, 1999). This can be

assessed using the aggregate metric in Equation 2.10:

$$\frac{(\mu_T - \mu_R)^2 + \sigma_T^2 - \sigma_R^2}{max(0.04, \sigma_R^2)}, \tag{2.10}$$

where $\sigma_T^2 = \sigma_{WT}^2 + \sigma_{BT}^2$ and $\sigma_R^2 = \sigma_{WR}^2 + \sigma_{BR}^2$ which tests the hypothesis in Equation 2.11,

$$H_0 : \nu_{PBE} = \sigma^2 + \sigma_T^2 - (1 + c_{FDA})\sigma_R^2 \geq 0 \quad for \quad \hat{\sigma_R^2} > 0.04$$
$$H_0 : \nu_{C.PBE} = \sigma^2 + \sigma_T^2 - \sigma_R^2 - (c_{FDA})0.04 \geq 0 \quad if \quad \hat{\sigma_R^2} > 0.04, \tag{2.11}$$

where σ_T^2 and σ_R^2 represent the subject variances of the T and R drugs, respectively. The aggregate statistic could be formulated using a mixed model approach based on a two period crossover design.

If the upper bound of the 90% CI for the BA parameters of interest (AUC or C_{max}) is below a pre-determined regulatory value of 1.74, PBE can be established for the metric under consideration.

2.4.4   Individual Bioequivalence Using Moment-Based Criteria

In order to declare that the T and R formulations are individually bioequivalent, the equation given below must be satisfied,

$$\Theta_{ind} = \frac{(\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{max(\sigma_{W0}^2, \sigma_{WR}^2)} < \theta_{ind}, \tag{2.12}$$

where $\sigma_D^2 = \sigma_{BT}^2 + \sigma_{BR}^2 - 2\rho\sigma_{BR}\sigma_{BT}$ and Θ_{ind} denotes the predetermined bound for the moment-based criterion for IBE assessment. However, Y_T, Y_R and $Y_R^{'}$ are no longer independent since only a replicate COD is used in IBE. Similarly,

$$E(Y_T - Y_R)^2 = var(Y_T - Y_R) + E[(Y_T - Y_R)]^2$$
$$= var(Y_T) + var(Y_R) - 2Cov(Y_T, Y_R) + (\mu_T - \mu_R)^2$$
$$= (\mu_T - \mu_R)^2 + \sigma_{BT}^2 + \sigma_{WT}^2 + \sigma_{BR}^2 + \sigma_{WR}^2 - 2\rho\sigma_{BR}\sigma_{BT}$$
$$= (\mu_T - \mu_R)^2 + \sigma_{WT}^2 + \sigma_{WR}^2 + \sigma_D^2,$$

and $E(Y_T - Y_R')^2 = 2\sigma_{WR}^2$.

Again, the FDA (1997) recommended 0.04 for σ_{W0}^2, 0.02 being the maximum difference for the within-variances of the T and R formulations, i.e $\sigma_{WT}^2 - \sigma_{WR}^2$, and a maximum value of 0.03 for interaction between subject and formulation, i.e σ_D^2. The value of σ_{W0}^2 is dependent on the individual difference ratio (IDR). The FDA guidance suggests the individual difference ratio (IDR) which is given as:

$$IDR = \sqrt{\frac{(\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 + \sigma_{WR}^2}{2\sigma_{WR}^2}}. \tag{2.13}$$

The largest tolerable value of the IDR is given as 1.25 by the FDA. Assuming the 80/125 rule with a certain probability, gives ln(1.25) as the upper bound for $\mu_T - \mu_R$. This results in the value of θ_{ind}:

$$\begin{aligned}\theta_{ind} &= \frac{average\ bioequivalence\ limit + variance\ allowance}{scaled\ variance} \\ &= \frac{[\ln(1.25)]^2 + 0.03 + 0.02}{0.04} \\ &= 2.49\end{aligned}$$

The IBE is a statistical method which is used to decide whether a patient being treated with the R drug can be switched to the T drug (Byron and Kenward, 2003). It takes into consideration the variability within each subject and the subject-by-formulation interaction. The aggregate metric is an approach used for the assessment of IBE (FDA, 1997):

$$\frac{(\mu_t - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{max(0.04, \sigma_{WR}^2)}, \tag{2.14}$$

which is used to test the following linearised null hypotheses:

$$H_0 : \nu_{IBE} = \sigma^2 + \sigma_D^2 + \sigma_{WT}^2 - (1 + c_{FDA})\sigma_{WR}^2 \geq 0 \ \ if \ \ \hat{\sigma_{WR}^2} > 0.04$$

$$H_0 : \nu_{C.IBE} = \sigma^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2 - 0.04(c_{FDA}) \geq 0 \ \ if \ \ \hat{\sigma_{WR}^2} \leq 0.04, \tag{2.15}$$

where σ_{WT}^2 and σ_{WR}^2 denote within-subject variability for the T and R drug formulations, respectively, σ_{BR}^2 and σ_{BT}^2 denote the variabilities between subjects for the R and T formulations, respectively, $\sigma_D^2 = \sigma_{BT}^2 + \sigma_{BR}^2 - 2\rho\sigma_{BT}\sigma_{BR}$ denotes the subject-by-formulation interaction with ρ being the between-subject

correlation of the T and R drugs, and μ_T and μ_R denote the means of T and R whilst $\delta = \mu_T - \mu_R$ denotes the difference between the means of the T and the R drugs.

The metrics in Equation (2.15) are scaled using the within variance $\hat{\sigma}^2_{WR}$ and a constant value of 0.04. Whenever the metric is scaled using $\hat{\sigma}^2_{WR}$, it is referred to as reference-scaled and it is called a constant scale when it is scaled using 0.04. The value of c_{FDA} is the regulatory value of 2.49 which is set by the FDA. The c_{FDA} assumes a within-subject variance for R of 0.04 and is calculated as follows:

$$
\begin{aligned}
c_{FDA} &= \frac{[\ln(1.25)]^2 + 0.03 + 0.02}{0.04}, \\
&= 2.49,
\end{aligned}
$$

which allows for a difference in means of $\ln(1.25)$ and a variance allowance of 0.03 for the subject-by-formulation interaction and an allowance of 0.02 for the difference in within-subject variances (FDA, 1997).

If the upper bound of the 90% CI for the BA parameters of interest, either AUC or C_{max}, is below a pre-determined regulatory value of 2.49, bioequivalence can be established for the metric under consideration.

2.5 Pharmacokinetic Parameters

The concentration time curve (CTC) is used to study the absorption and e-limination rate of a drug in a blood sample in a comparative BA study (Chow and Liu, 2008). At various time points, blood samples are taken after a drug is administered. There are other pharmacokinetic metrics but they only provide support information and are not used to approve the new formulation. It is argued that AUC and C_{max} are insufficient to provide evidence that a T drug is bioequivalent to the R drug; however, regulatory bodies have depended on and used the AUC and C_{max} as the main metrics for bioequivalence approval of drugs (Lacey *et al.*, 1995; Steinijans *et al.*, 1995; Rescigno and Powers, 1998). AUC is a PK parameter which determines the magnitude of absorp-

tion of a drug, thus, the total quantity of the drug that is absorbed by the body.

The AUC is regarded as the best standard measure for extent of BA while the C_{max} as a measure of the rate of BA is criticized for not characterizing the rate of BA appropriately (Cartwright, 1991; Herchuelz, 1996; El-Tahtawy *et al.*, 1998). The C_{max} is heavily dependent on the sampling scheme and has more variability than the AUC and therefore poses problems for assessment of bioequivalence (Buice *et al.*, 1996; Tsang *et al.*, 1996; Patterson, 2003). Despite these shortfalls of the C_{max}, it is still used as one of the metrics of bioequivalence studies because it performs better than other measures (Bois *et al.*, 1994). During bioequivalence studies, each of these endpoints are analysed separately and are generally log-transformed (Westlake, 1979; Midha *et al.*, 1993). The normality of the AUC and C_{max} are not tested after log transformation, as stated in FDA (1992) and statistical analysis is also not conducted on the original metrics even if there is evidence that the data are not normally distributed.

During the sampling stage, the plan outlined for taking the blood samples should be designed properly for suitable evaluation of the pharmacokinetic parameters for both rate and absorption. These pharmacokinetic parameters are mostly determined either directly (model-based approach) or indirectly (concentration-time profile approach). An example of the concentration - time graph approach is illustrated in Figures 2.5 and 2.6. Methods for estimating the AUC as suggested by Chow and Liu (2008) include linear interpolation which uses the trapezium rule, a planimeter, spline and lagrange methods and a physical method. However, the linear interpolation technique is the most widely used. For the linear interpolation method using the trapezoidal rule, let $C_0, C_1, ...C_k$ be the blood concentrations obtained at times $0, t_1, ...t_k$. Then AUC_{0-t_k} is obtained using Equation 2.16,

$$\text{AUC}_{0-t_k} = \sum_{i=2}^{k} \left(\frac{C_{i-1} + C_i}{2} \right) (t_i - t_{i-1}). \tag{2.16}$$

Although the AUC is normally calculated from time zero to time t, Hauschke *et al.,* (2007), suggested that AUC from zero to infinity in a single-dose study should be determined. Martinez and Jackson (1991) articulated that a fraction of area under t_k to infinity may be large, should the level of blood at t_k be significant. The AUC from time zero to infinity, which is represented as $AUC_{0-\infty}$, is approximated using the equation shown in Equation 2.17 (Rowland, 1980; Chow and Liu, 2008),

$$\text{AUC}_{0-\infty} = AUC_{0-t_k} + (\frac{C_k}{\lambda}). \tag{2.17}$$

This serves as the extent of absorption in single-dose studies. After the drug is administered, C_k is the concentration of blood which occurs at the last sampling time point, λ is the elimination rate constant and given as $\lambda = -2.303 \times$ gradient of the terminal segment of the log concentration time curve (Chow and Liu, 2008). The time interval for the calculation of each λ as well as $AUC_{0-t}/AUC_{0-\infty}$ is indicated in Table 2.1. According to Hauschke *et al.* (2007), the percentage ratio of $AUC_{0-t}/AUC_{0-\infty}$, as shown in Table 2.1, should exceed 80% for each subject. This is important not to over extrapolate the $AUC_{0-\infty}$. From Table 2.1, only one subject had the percentage ratio to be 78% while the rest are above 80%. The values for the percentage ratio, $AUC_{0-t}/AUC_{0-\infty}$, therefore confirms that the choice of the ultimate section of the logarithm concentration time curve used to calculate the lambda value is consistent and accurate.

Table 2.1: Information used for the calculation of $AUC_{0-\infty}$ which includes the estimate of λ for the terminal rate constant and the extrapolated AUC fraction.

Subject	Sequence	Interval for estimating λ (h)	AUC_{0-t}	λ	C_t	$AUC_{0-\infty}$	$\frac{AUC_{0-t}}{AUC_{0-\infty}}$	$t_{1/2}$
1	RT	4-72	53.378	0.031385	0.168	58.7309	0.91	22.09
2	RT	6-72	63.4214	0.04476	0.0912	65.4589	0.97	15.49
3	RT	6-72	78.4096	0.052737	0.0883	80.08395	0.98	13.14
4	RT	8-72	56.0545	0.063779	0.176	58.8140	0.95	10.87
5	RT	6-72	38.5983	0.029050	0.170	44.4499	0.87	23.86
6	RT	12-72	78.99	0.037320	0.179	83.78636	0.94	18.57
7	RT	8-72	71.838	0.03518	0.154	76.21549	0.94	19.70
8	RT	4-48	54.3085	0.030158	0.463	69.66098	0.78	22.98
9	TR	4-30	56.264	0.12395	0.151	57.4822	0.98	5.59
10	TR	6-48	61.1015	0.071681	0.112	62.6640	0.98	9.67
11	TR	8-72	59.2165	0.02562	0.275	69.9503	0.85	27.05
12	TR	4-48	60.798	0.05036	0.330	67.3508	0.90	13.76
13	TR	3-30	68.9234	0.18584	0.0513	69.1994	0.99	3.73
14	TR	3-24	67.7175	0.21020	0.105	68.2170	0.99	3.30
15	TR	6-72	66.153	0.034855	0.177	71.2319	0.93	19.89
16	TR	4-48	59.9715	0.03546	0.445	72.5209	0.83	19.55

However, there are challenges in estimating the terminal portion of the log-linear plot for the concentration-time graph as indicated in the concentration time graph (Figures 2.5 and 2.6).
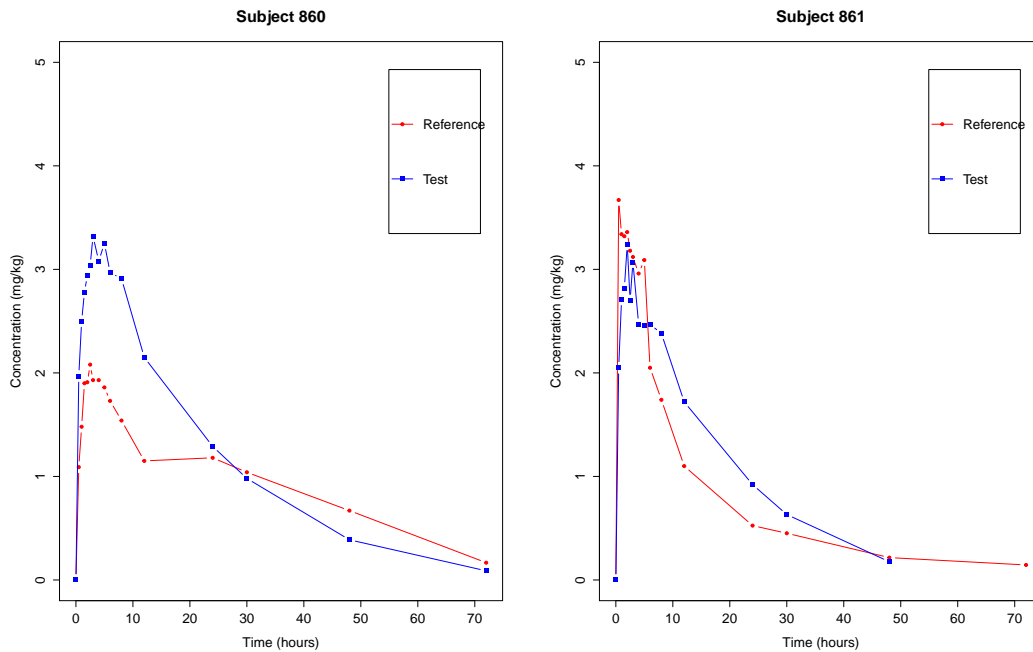
Figure 2.5: Concentration profile for T and R formulations of subject 860 in sequence 1.

Figure 2.6: Concentration profile for T and R formulations of subject 861 in sequence 1.

The choice of the range of the terminal portion affects the estimation of the extrapolated AUC to infinity. A very suitable range should be selected either using a mathematical algorithm or by a clinical pharmacokineticist together with the person who determined the concentrations of the subjects.

$$f.dose = clearance.AUC_{0-\infty}, \ 0 < f \leq 1, \tag{2.18}$$

From Equation 2.18, it can be established that the fraction, f, of a dose absorbed is proportional to the $AUC_{0-\infty}$ where clearance is a proportionality factor. Clearance, as defined by Rowland and Tozer (1995), is the apparent volume of blood that is completely removed of drug per unit of time.

For a compartment model in Equation 2.19 with first-order absorption and elimination kinetic, the rate of constant of absorption is estimated with the residuals method (Gibaldi and Perrier, 1982),

$$C_t = \left(\frac{K_a F D_o}{V(K_a - K_e)} \right) \left(e^{-K_e t} - e^{-K_a t} \right), \tag{2.19}$$

where C_t represents the concentration, K_a represents the rate of constant of absorption, K_e represents the rate of constant of elimination, D_o represents the quantity of the dose that is administered, F represents the proportion of the dose that is absorbed or gets to the systemic circulation and V is the distribution volume. C_{max}, which is the maximum concentration, is defined mathematically as $\hat{C}_{max} = max(c_0, c_1, ..., c_k)$. T_{max}, in Equation 2.20, is the time needed to attain maximum concentration and is defined as the estimated time at which C_{max}, in Equation 2.21, is observed.

$$T_{max} = \frac{2.303}{(K_a - K_e)} \log \left(\frac{K_a}{K_e} \right) \tag{2.20}$$

$$C_{max} = \left(\frac{K_a F D_o}{V(K_a - K_e)} \right) \left(e^{-K_e t_{max}} - e^{-K_a t_{max}} \right) \tag{2.21}$$

The half-life $(t_{1/2})$ and the elimination rate constant (K_e) are often the parameters which are studied during the elimination phase of a drug, according to Chen and Pelsor (1991) and Chow and Liu (2008). The elimination half-life of a drug is explained by Chow and Liu (2008) as the time which it takes for the concentration of the drug to reduce by one-half from the body. The $t_{1/2}$ can be derived with the assumption that the reduction in concentration of blood is of first order, as shown in Equations 2.22, 2.23 and 2.24,

$$\log(D) = \log(D_o) - \frac{K_e t}{2.303}, \tag{2.22}$$

where D represents the quantity of the drug in the body and is given as $D = \dfrac{D_o}{2}$ $(t = t_{\frac{1}{2}})$,

$$\log \left(\frac{1}{2} \right) = \frac{-K_e t_{\frac{1}{2}}}{2.303}, \tag{2.23}$$

39

$$t_{\frac{1}{2}} = \frac{0.693}{K_e}, \qquad (2.24)$$

where $K_e = -2.303 \left(\frac{d\log D}{dt} \right)$.

Knowledge of half-life is necessary to determine the frequency of administration of a drug so as to get the desired plasma concentration. It is independent of the dose administered.

2.6 Power of a Test

When planning bioavailability analysis, it is very important to choose the right sample size as this affects the power of the study (Wang and Bakhai, 2006). The sample size, according to bioequivalence studies, is the total number of subjects involved in the study. This number, as stated by Hauschke *et al.* (2007), is determined by the amount of variability in the PK characteristic, the power of the test, the level of significance and the expected deviation of the T from the R formulation. A larger sample size has a better power and has the ability to detect any treatment effect while a small sample size has less power and could necessarily not be able to accurately detect a treatment effect. That is, by taking larger samples, the ability to find a difference in means of the two formulations, if they do exist, is improved. Also, the power of the test decreases as the population variance reduces and as the difference in the means increases, the power also increases.

The power of the test, according to Chow and Liu (2008) and Hauschke *et al.* (2007), is defined as the likelihood of rejecting the null hypothesis of bioinequivalence between an R and a T drug, while the alternative hypothesis of bioequivalence is true.

$$H_0 : bioinequivalence$$
$$H_1 : bioequivalence. \qquad (2.25)$$

40

This implies that the power of a test is the likelihood of concluding correctly that a drug is effective and bioequivalent when it is. According to Owen (1965), power in BE studies is the likelihood of demonstrating that two drugs are bioequivalent correctly when the two drug formulations are indeed bioequivalent.

$$power = 1 - \beta,$$

$$= P(reject \ H_0 \ when \ H_0 \ is \ false),$$

where $\beta = P(type \ II \ error)$.

The power $(1 - \beta)$ of the decision rule is the chance of correctly concluding bioequivalence.

There are two types of errors that depend on the sample size and the power of the test. The two types of errors are used when testing the null hypothesis against the alternative hypothesis for average bioequivalence. In ABE, there are chances of incorrectly establishing that two drug formulations are bioequivalent when they are not. This phenomenom is referred to as making a Type I error. A Type I error is perpetrated whenever the null hypothesis (bioinequivalent) is rejected whilst it is indeed true and a Type II error occurs when the null hypothesis is false but it is not rejected. The two types of errors are shown in Table 2.2 (Chow and Liu, 2008).

Table 2.2: Type I and Type II errors for a hypothesis test.

	The null hypothesis is	
	True	False
Fail to reject null hypothesis	Correct decision	Type II error β
Reject null hypothesis	Type I error α	Correct decision

Usually the sample size used in bioequivalence studies, as stated by Chow and Liu (2008), is chosen based on a power function which tests the null hypothesis

of bioequivalence between the two formulations ($\mu_T = \mu_R$). In every BE study it is important to choose the sample size while considering the Type I error rate as well as the equivalence criteria, the power (normally 90%) and the intra-subject variation (Patterson, 2003). Whenever the intra-subject coefficient of variation (CV) increases beyond 30%, the sample size required for BE increases as shown in Table 2.3 (Patterson, 2003).

Table 2.3: Sample sizes producing 90% power in BE for Two period crossover design(COD).

CVw%	Two period COD^*	Two period COD
30	40	50
	54	60
	112	124
45	84	90
	112	120
	230	244
60	140	146
	184	194
	384	404
75	200	206
	264	276
	554	574

Two period COD^* assumes subject-by-formulation interaction is negligible. Two period COD assumes subject-by-formulation interaction is non-negligible.

Patterson (2003) indicated that the number of samples needed for BE studies reduces by half when using replicate designs with high intra-subject variability. However, when the variability is low, a non-replicate two period COD is preferred as the replicate design does not improve precision drastically.

In order to test for bioequivalence between an R and a T drug, two types of hypothesis approaches are used, namely direct and indirect bioequivalence testing (Hauschke *et al.*, 2007). The indirect bioequivalence testing uses the hypothesis in Equation 2.26,

$$H_0 : bioequivalence$$
$$H_1 : bioinequivalence.$$

(2.26)

The major shortcoming of this indirect approach, according to Hauschke *et al.* (2007), is the fact that the chances of judging bioinequivalence mistakenly is regulated and therefore is not accepted by the regulatory authorities.

Table 2.4: Type I and Type II errors for indirect bioequivalence testing.

	The null hypothesis of bioequivalence is	
	True	False
Fail to reject null hypothesis of bioequivalence	Correct decision	Consumer risk β
Reject null hypothesis of bioequivalence	Producer risk α	Correct decision

The ultimate regulatory concern for regulatory bodies is to control consumer risks, thus limiting the chances of concluding incorrectly bioequivalence (Hauschke *et al.*, 2007). Due to this, the direct approach is the approved method by the regulatory bodies as in Equation 2.25.

The producer and consumer risks for the direct method of BE assessment are illustrated in Table 2.5.

Table 2.5: Type I and Type II errors for direct bioequivalence testing.

	The null hypothesis of bioinequivalence is	
	True	False
Fail to reject null hypothesis of bioinequivalence	Correct decision	Producer risk β
Reject null hypothesis of bioinequivalence	Consumer risk α	Correct decision

In ABE, Type I error is defined as the likelihood of deciding that two drug formulations are bioequivalent. This probability is also known as alpha, confidence or regulatory risk.

The relationship between a Type I and a Type II error for a bioequivalence study, as illustrated by Chow and Liu (2008), is shown in Table 2.6. The power is determined based on a hypothesis and the outcome. There are four possible outcomes with hypothesis testing, two of which are correct decisions while the other two are incorrect, as shown in Table 2.6. The inferences which are incorrect are the errors. The two types of errors, Type I and Type II, in some cases, are referred to as consumer's risk and producer's risk, respectively, as explained in Table 2.6 (Patterson, 2003).

Table 2.6: Type I and Type II errors.

	True State H_0	
	Bioinequivalent	Bioequivalent
Bioinequivalent	Right decision	Type II error
Bioequivalent	Type I error	Right decision

The power of the test is not the main issue of this study but rather the comparison of the formulation means and shall not be investigated further.

2.7 Simulated Annealing Algorithm

Many problems in computer science, engineering and manufacturing can be modelled as minimisation or maximisation of a cost function or objective function over a finite set of discrete variables (Aarts and Korst, 1990). This class of combinatorial optimisation problems has received attention over decades and major achievements have been made in its analysis (Papadimitriou and Steiglitz, 1998). Solving such combinatorial optimisation problems amount to finding the optimal solution among a finite number of alternative solutions. According to Aarts, Korst and Michiels (2005), combinatorial optimisation problems are separated into subclasses. The first class contains problems that can be solved efficiently using known algorithms (linear programming) and the second class contains problems that are difficult to solve and are formally referred to as nondeterministic polynomial time (NP) hard. There is no known algorithm that gives an optimal solution for an NP-hard problem. Many combinatorial optimisation problems belong to the NP-hard group. The travelling salesman problem (TSP) is one of the well known combinatorial optimisation problems. In a TSP, a salesman starts from his home city and must visit each city once on a prescribed list of cities to be visited and return home while minimising the tour length and total cost of travelling (Aarts and Korst, 1990). Solving combinatorial optimisation problems aim at finding the minimum or maximum value of a function (cost or objective function).

Annealing is the physical process of heating up a solid until it melts followed by cooling it down until it crystallises into a state with a perfect lattice (Aarts and Korst, 1990). The atoms in the material have high energies at high temperatures and have freedom to restructure themselves. The atomic energies decrease until a state of minimum energy is obtained. Whenever the cooling process is rapid (quenching), it results in defects in the crystal structure (Pham and Karaboga, 2012). According to Aarts and Korst (1990), the cooling process must be done cautiously in order not to get trapped in a local optimal

solution. This process can be formulated in a combinatorial optimisation as a problem of finding a solution with minimal cost among a potentially large number of solutions. By creating a link between the cost function and the free energy, and between the solutions and the physical states of the atoms, a solution method in combinatorial optimisation based on the simulation of the physical annealing process is used (Aarts and Korst, 1990). This method is known as Simulated Annealing. The annealing concept in combinatorial optimisation was introduced in the early 1980's (Kirkpatrick, Gelatt and Vecchi ,1983).

2.7.1   Metropolis Algorithm

Annealing is a thermal process for obtaining low energy states of a solid in a heat bath (Aarts and Korst, 1990). The annealing process consists of the following two steps (Kirkpatrick *et al.* 1983):

1. increase the temperature of the heat bath to a maximum value at which the solid melts; and

2. decrease carefully the temperature of the heat bath until the particles arrange themselves in the ground state of the solid.

Metropolis *et al.* (1953) introduced a simple algorithm that can be used to provide an efficient simulation of a collection of atoms in equilibrium at a given temperature. The algorithm is based on Monte Carlo techniques and generates a sequence of states of the solid. Given a current state i with energy E_i, a subsequent state j is generated by applying a perturbation mechanism which transforms the current state into the next state. The energy of the state j is denoted by E_j. If the difference in energy, $E_j - E_i \leq 0$, the state j is accepted as the current state. However, if $E_j - E_i > 0$, the state j is accepted with a probability which is given as

$$exp(\frac{E_i - E_j}{k_B T}),$$

where T represents the temperature of the heat bath and k_B is a physical constant known as the Boltzmann constant. This acceptance rule is known as the Metropolis criterion. When the temperature is sufficiently slowly lowered, the solid can reach thermal equilibrium at each temperature (Aarts and Korst, 1990). This process is achieved in the Metropolis algorithm (MA) by generating a large number of transitions at a given temperature. The Boltzmann distribution, which is a characteristic of thermal equilibrium, gives the probability of the solid being in a state i with energy E_i at temperature T, and is represented by

$$P_T\{X = 1\} = \frac{1}{Z(T)} exp \left(\frac{-E_i}{k_B T} \right),$$

where X denotes a stochastic variable representing the current state of the solid and Z(T) represents the partition function which is defined as

$$Z(T) = \Sigma_j exp \left(\frac{-E_j}{k_B T} \right),$$

where the summation extends over all possible states. The Boltzmann distribution plays an important role in the SA algorithm analysis.

2.7.2   Simulated Annealing

The MA can be applied to create a sequence of solutions of a combinatorial optimisation problem by assuming the following equivalences between a physical many-particle system and a combinatorial optimisation problem (Aarts and Korst, 1990):

1. solutions in a combinatorial optimisation problem are equivalent to states of the physical system; and

2. the cost of a solution is equivalent to the energy of a state.

Metropolis *et al.* (1953) suggested the MA for simulation of a solid in a heat bath to simulate thermal equilibrium based on the Monte Carlo approach. Each state has a distinct level of energy which is defined by the objective

function. A difference between the new state and old state, $\Delta E = E_j - E_i$, is calculated and the new state is chosen with a likelihood of $e^{\frac{-\Delta E}{K_T}}$. However, should the new state have a lower energy than the current state ($\Delta E < 0$), the new state is accepted as the starting point with a likelihood that is dependent on the energy level difference as well as the current temperature shown in Equation 2.27,

$$P(accept) = \begin{cases} 1 & \text{if } \Delta E \leq 0, \\ e^{\frac{-\Delta E}{K_T}} & \text{if } \Delta E > 0, \end{cases} \qquad (2.27)$$

where P represents the acceptance probability and T denotes the control parameter. The algorithm has the ability to escape the local minimum based on the acceptance of the worst state. This acceptance is referred to as the Metropolis criterion (Metropolis *et al.*, 1953).

MA was applied in optimisation problems by Kirkpatrick *et al.* (1983) who named it the Simulated Annealing (SA). SA is a stochastic global optimization method which was described first by Metropolis *et al.* (1953) who developed a Metropolis algorithm to simulate a collection of atoms at a given temperature. Kirkpatrick et al. (1983) showed how the model for simulating the annealing of solids, proposed by Metropolis *et al.* (1953), can be used for solving optimisation problems where the objective function to be optimised is similar to the energy level of a solid. SA consists of two nested loops. The decrement of the temperature from an initial high to a final low temperature is the outermost loop which is also known as the cooling schedule. The innermost loop repeats the process over a specified Monte Carlo simulation. The SA algorithm initially begins with a high temperature where so many states are accepted. However, as the temperature decreases, only few states are accepted. The chances of accepting a bad state is done by comparing it with a random number generated uniformly from [0,1]. The temperature is reduced when the Metropolis iteration loop is completed and the process is repeated until the stopping criterion is met.

The MA starts from an initial energy level, x, and the system is disturbed at random to a new energy level, y, in the neighbourhood of x. The value of the objective function f_y is calculated. If the change, $\Delta f_{xy} = f_y - f_x$, is a reduction in the objective function value, the new state is accepted. However, should the change be an increase in the value of the objective function, the new state is accepted with a likelihood $\frac{\Delta f_{xy}}{T}$ with T being a control parameter.

The SA algorithm is seen as an iteration of MA which is evaluated at decreasing values of the control parameter that takes on the role of temperature. The main attributes of the SA algorithm are its general applicability (or usefulness) and ability to obtain solutions without being trapped in a local minimum (Aarts and Korst, 1990). This is done by not only accepting better solutions but also worse solutions with a given probability. The SA algorithm has the ability to find high-quality solutions which do not necessarily depend on the choice of the initial solution and therefore is regarded as a very effective and robust algorithm (Aarts and Korst, 1990).

The flowchart of SA is given in Figure 2.7.
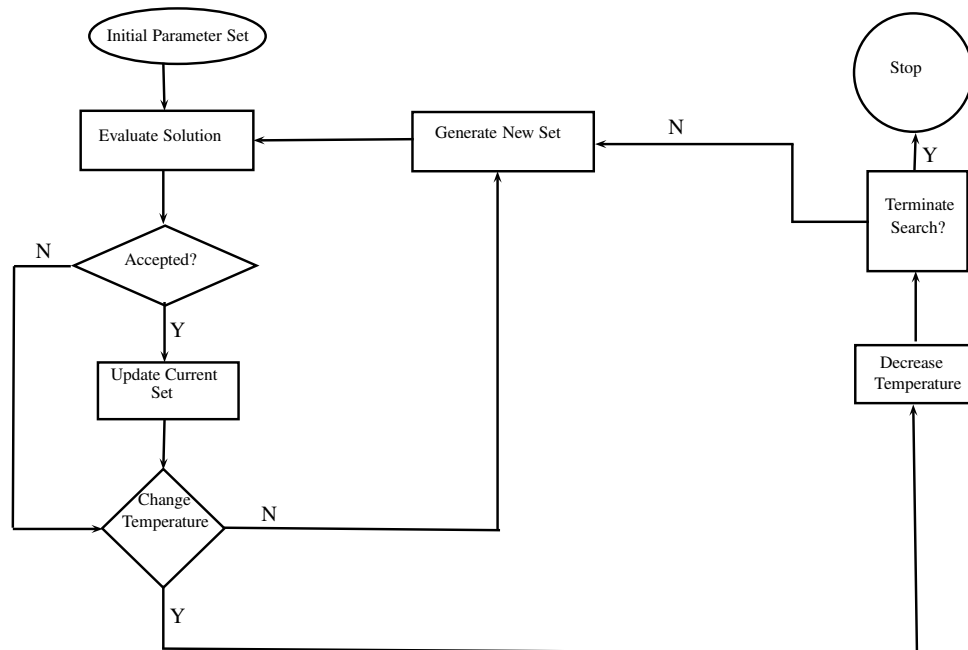
Figure 2.7: The SA Flowchart.

If the current solution (f_{new}) has the value of the objective function lower compared to that of the previous solution (f_{old}), the current solution is accepted (in a minimisation situation). The current solution may be accepted if the Boltzmann distribution in Equation 2.28,

$$e^{\frac{-f_{new}-f_{old}}{T}},\tag{2.28}$$

is greater than a uniform random number in $[0,1]$, where T is the temperature control parameter.

The annealing process consists of a number of steps which are described in the section below.

2.7.2.1 Initial Population

At each iterative stage a definition of an initial guess for the parameter values is required. SA does not require the use of several initial solutions. One approach is to randomly select the initial parameter values given a set of appropriate boundaries. The closer the initial estimate is to the global optimum, the quicker the optimisation process.

2.7.2.2 Initial Temperature

The control parameter, T, should be defined with care as it controls the acceptance rule defined by Equation 2.28. T should be large to avoid being trapped in a local minimum but small enough to move off a global minimum.

2.7.2.3 Perturbation Mechanism

This mechanism creates a new solution from the current solution by exploring the neighbourhood of the current solution and making small changes to the current solution. When the parameters are continuous variables, a solution P is defined as a vector $\boldsymbol{P} = (x_1, x_2, ..., x_n)$ which represents a point in the search space. A new solution is generated using a vector $\boldsymbol{\varphi} = (\varphi_1, .., \varphi_n)$ of standard deviations to create perturbation from the current solution. A neighbourhood solution is produced from the current solution by $x_{i+1} = x_i +$

$N(0, \sigma_i^2)$ where $N(0, \sigma_i)$ is a random Gaussian number with mean of zero and standard deviation σ_i.

2.7.2.4 Objective Function

The objective function or the cost function relates to the parameter that needs to be minimised or maximised. For example, there may be the need to minimise the number of sampling times which eventually reduces the cost.

2.7.2.5 Initial Value of Temperature

A basic assumption which underlies the initial value calculations of the control parameter T (temperature) is that it should be large for the movement between different energy states to be accepted at this value. It is usually achieved through generating trials m_0 and requires that the initial acceptance ratio, $x = x(T_0)$, is close to one, where $x(T)$ is the ratio between the number of accepted transitions and the number of proposed transitions. The initial value of T is obtained from the expression:

$$T_0 = \overline{\Delta f^+} \left(ln \frac{m_2}{m_2 x_0 + (1 - x_0)m_1} \right)^{-1}, \qquad (2.29)$$

where m_1 and m_2 denote the number of trials $(m_1 + m_2 = m_0)$ with $\Delta f_{xy} \leq 0$ and $\Delta f_{xy} > 0$ and, $\overline{\Delta f^+}$ the average value of Δf_{xy} values for which $\Delta f_{xy} > 0$, and x_0 represents the initial configuration.

2.7.2.6 Decrement of the Control Parameter

The new value T_{t+1}, which is calculated from T_t, is obtained from the expression

$$T_{t+1} = T_t (1 + \frac{T_t ln(1 + \delta)}{3\sigma(T_t)}), \qquad (2.30)$$

where $\sigma(T_t)$ is the standard deviation of the cost function values at the points in the Markov chain (MC) at T_t. The constant δ is the distance parameter and determines the rate of decrease of the control parameter.

2.7.2.7 Stopping Criterion

This is based on the idea that the average function value $\overline{f(T_t)}$ over an MC decreases with T_t so that $\overline{f(T_t)}$ converges to the optimal solution as $T_t \to$ approaches 0. The MC stops whenever there are no changes in $\overline{f(T_t)}$. The SA algorithm is terminated if

$$\left| \frac{d\bar{f}_s(T_t)}{dT_t} \frac{T_t}{\bar{f}(T_0)} \right| < \varepsilon_s, \tag{2.31}$$

where $\bar{f}(T_0)$ is the mean value of f at the points in the initial Markov chain, $\bar{f}_s(T_t)$ is the smoothed value of $\bar{f}$ over a number of chains so as to reduce the fluctuation of $\bar{f}(T_t)$, and ε is a small positive number called the stopping parameter.

The proposed SA algorithm steps used in this dissertation is adopted from Choi *et al.* (2007):

1. The initial temperature T_0, and the number of MC simulations at each temperature P is set;

2. The initial sampling times $t_0 = (t_1, t_2, ..., t_n)$ are set;

3. Generate the new sampling schedule t_{new};

4. Estimate the energy of the current solution $(E_{current})$ and new schedule (E_{new});

5. Check if the new schedule should be accepted using the acceptance criterion;

6. Repeat the process from steps 3 to 5 for p number of times;

7. Calculate next temperature; and

8. Repeat the process until the stopping criteria is effected.

2.8 Population PK

In a traditional non-compartmental approach of bioequivalence, the subjects are recruited from a homogenous population which makes it difficult to make an inference about the whole population. The solution to this is through population pharmacokinetics (PopPK). This approach uses a nonlinear mixed effects model by building a PK model with fixed effects, between-subject variabilities and residual variabilities. The use of PopPK modelling has several advantages such as investigating the characteristics of the PK parameters in a model (Chow and Liu, 2008). The advantages and disadvantages of the PopPK method, as stated by Bonate and Steimer (2006), are summarised in Table 2.7.

Table 2.7: Advantages and disadvantages of PopPK method.

Advantages	Disadvantages
Subgroups within the population can be distinguished which may not have been noticed.	It is very expensive.
Population method can analyse thinly dispersed data.	It is time consuming.
Essential covariates are used and may be able to interpret subject variability.	The methods are not easy to comprehend.
The information that are obtained could be used for prediction of dose for individual subjects.	Different PopPK analysts may develop different models.
It is a model based on promoting the use of prior knowledge, thereby enhancing understanding and statistical power.	Difficult to review.

PopPK, as defined by Aarons (1991), studies the origins and the variation in the concentration of drugs in individual subjects who are sampled from a population and receive a quantity of the drug of interest at a particular time period. That is, population PK studies the connection that exists between the quantity of administered dose and concentration of the sampled subjects as well as the population from which they were sampled. The population PK parameters are evaluated by fitting a model. There are a number of methods used to estimate the parameters of the PopPK including the traditional approach which uses the standard two-stage (STS) method and the nonlinear mixed effects modelling approach (NLMEM) (Chow and Liu, 2008). These two methods are discussed in the next section. In this dissertation, the NLMEM approach is adopted because of its advantages that will be discussed in Section 2.10.

2.8.1 Traditional Two-stage

This approach of estimating the PopPK parameters has two stages, hence the name: standard two-stage (STS) approach. The PopPK parameters are estimated by fitting each individual subject's data separately and later combining the individual estimates (Chow and Liu, 2008).

For each subject, enough dosage is administered and sufficient blood concentrations are drawn so that the PK parameters are approximated accurately for each subject. The estimates of the parameters are acquired using a deterministic PK model which could be a one or multi compartment model (Chow and Liu, 2008).

From the PK parameter estimates which were derived individually from each subject in the first stage, covariance model analysis is considered to avoid confounding effects as well as the interactions that may exist among the covariates (demographics) and to study the treatment effect (dosage and route of admin-

istration) and the variation between individuals (Chow and Liu, 2008).

However, this approach has cost and ethical issues as stated by Sheiner *et al.* (1977) and also has disadvantages from a statistical point of view. The variation of each individual is not considered whenever the estimates obtained from the deterministic model in the first stage are used. It becomes difficult to obtain reliable and accurate approximations of the PK parameters for each subject whenever the sampling of the subjects is sporadic (Chow and Liu, 2008). The traditional approach is incapable of describing the attributes of the population where a number of demographic-, physiological- and behavioural attributes are documented for each individual subject including their age, weight, sex, and other variables (Chow and Liu, 2008).

2.8.2 Nonlinear Mixed Effects Modelling Approach

The traditional two-stage method described in section 2.9 has several setbacks and therefore an alternative method, the NLMEM approach, is proposed and comes highly recommended by the FDA. This method is a nonlinear regression model which accounts for random and fixed effects. NLME models are quite similar to the linear mixed effect models (LMEM) but the function in NLMEM is nonlinear (Bonate and Steimer, 2006). PopPK studies the pharmacokinetics in a population and it models the data from every individual subject concurrently, but it is also able to account for within-subject variability, between-subject variability and residual variability (Bonate and Steimer, 2006). This method was introduced by Sheiner and Beal (1985) and is the most widely used method for population studies. During NLMEM the attributes of the population constitute the mean of the population which is obtained from fixed effect parameters, the variation, which is within the population, as well as the variability from the random effect parameters (Bonate and Steimer, 2006).

One objective of a PopPK is to develop a model which connects the concentra-

tion to the administered dose and other covariates as well as to obtain estimates of the PK parameters (Bonate and Steimer, 2006). These NLME parameters are derived using a maximum likelihood estimation (MLE) method. Because of the nonlinear dependence of the observations it becomes demanding to calculate the likelihood of the data. Hence, two algorithms are widely used: the first-order (FO) approximation and the stochastic approximation expectation-maximisation (SAEM). The NLME does not require the data to be necessarily frequent and does not follow a structured sampling time schedule (Bonate and Steimer, 2006).

2.8.3 Pharmacokinetic Models

The use of PK modelling in drug development has received much attention (Sheiner and Steimer, 2000). Teorell (1937) first introduced the PK models to describe drug absorption, distribution and elimination. He suggested that organs and tissues be separated into different compartments and are only connected by kinetic rate. In this dissertation, the focus is on compartmental models even though there are different types of models, including physiological models and empirical models, as stated by Shargel *et al.* (2004).

Consider a single one-compartment model which is a bolus (administration of a discrete quantity of a drug so as to raise its concentration to an effective level) intravenous injection (IV), as described by Welling (1997), and given in Figure 2.8.

$$D \xrightarrow{\ K_a\ } \boxed{A(t)} \xrightarrow{\ K_e\ }$$
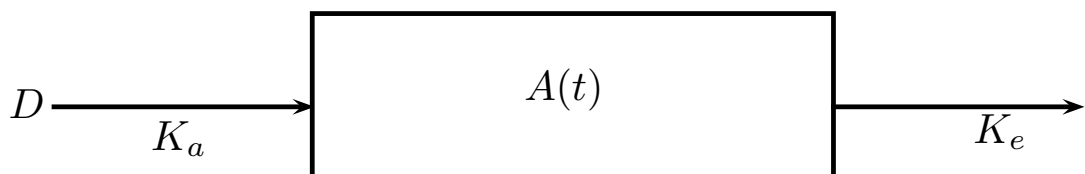
Figure 2.8: One compartment model with bolus IV.

Here, A represents the amount of the drug in the body calculated as $A = C \times V$, where C is the concentration of the drug in the body and V is the volume of distribution. K_e represents the elimination rate constant and K_a represents the absorption rate constant. Using a differential equation, the following equation is obtained:

$$\frac{dA}{dt} = -K_e A, \tag{2.32}$$

with t being the time. After performing integration Equation 2.32 becomes:

$$C = C_0 e^{-K_e t}. \tag{2.33}$$

$$lnA - lnA_0 = -K_e A$$

$$e^{ln(\frac{A}{A_0})} = e^{-K_e t}$$

$$A = A_0 e^{-K_e t}$$

The amount of drug A can be expressed as the concentration simply by dividing it by the volume of distribution, V, of the drug.

2.9 Summary

Bioequivalence studies started in the early 1960s to 1970s to meet the needs of consumers in having access to inexpensive and efficacious drug products. The origin of bioequivalence studies was purely for economic reasons and was driven by legislation across different countries. Therapeutic failures prompted extensive research in the area of bioequivalence studies and led to the establishment of techniques and regulations for assessing whether two drug products were bioequivalent based on the extent of BA between different formulations.

ABE was then introduced as a measure of assessing two drug products and was adopted by the FDA and other bodies such as SAMCC. ABE, although it protected the health of the public since it was adopted, could only compare the

mean rate and extent of BA and does not compare between- or within-subject variability between the drug formulations. This led to further research on PBE and IBE.

The literature presented in this Chapter shows that research on bioequivalence studies employed ABE, PBE and IBE to show that two drug products were indeed bioequivalent. However, very little is done using IBE since that requires a higher order COD and most of the designs used are limited to 2×2 COD.

Population PK studies is one area of research which is gradually coming to the fore. This method focuses on investigating the population characteristics of the PK parameters using a modelling technique. Available literature indicates that different methods, such as the traditional approach, nonlinear mixed effect modelling technique and others, have been used to model the population parameters. However, nonlinear mixed effect modelling is widely used and recommended by the FDA.

The techniques which were discussed in this section form the basis of the methodology which is discussed in the next Chapter and used in the dissertation.

Chapter 3

Methodology

3.1 Introduction

For a T drug to be approved as a replacement for the R drug and be marketed, regulatory bodies, including the FDA, WHO and SAMCC, demand that ABE between the T and R drugs should be provided (FDA 2001). BE is measured using C_{max} and AUC which measure the extent and the rate of drug absorption, correspondingly. The scope of this research is to perform statistical analysis on the PK parameters using the available data, estimate the population- and individual absorption rate, clearance and volume of distribution of the drug using a model and also optimise (minimise) the samples of blood taken from each subject. The 2×2 crossover design will be used for this dissertation as well as the TOST. ANOVA will also be carried out and the results will be used in the CI approach. A confidence interval approach will be used to determine bioequivalence using the results obtained from the ANOVA analysis. An SA algorithm will be used to optimise the blood sampling times taken from the subjects.

Bioequivalence assessment between the T and R drugs can be assessed through four main ways (FDA, 2003):

- In vitro studies;

- Pharmacodynamic studies;

- Comparative clinical studies; and

- Pharmacokinetic studies.

These methods can be used singly or as a combination depending on the type of treatment formulation or the method of administration of the drugs. According to the FDA, pharmacokinetic studies are the preferred method for bioequivalence studies. This choice of bioequivalence studies as the preferred method for bioequivalence studies was supported by Wang and Bakhai (2006). Pharmacodynamic studies and comparative studies are most recommended whenever pharmacokinetic studies cannot be applied. In vitro studies are generally recommended for immediate release (IR) of drugs that are taken orally and are in a solid form (FDA, 2003).

3.2   Experimental Design

Before BA studies are conducted the type of design to be used, subject selection criteria, the dosing schedules as well as the statistical methods to be used, are considered (Chow and Liu, 2008). The types of statistical analyses to be performed depend on the type of design that is used. For example, when one wants to conduct IBE, the most appropriate design according to the FDA (2001), is the repeated measures design (preferably two sequence, four period design) as this enables easy computations of the within and between subject variabilities and the subject-formulation relations. In general, meaningful statistical inferences can be drawn from data only when appropriate designs are used.

When ABE and PBE are the only chosen methods for BE comparisons, a non-replicate standard two-formulation, two-period, two-sequence COD is used since using a replicated design will be more expensive and time consuming (FDA, 2001). During BE studies the study should be properly designed so as

to distinguish the drug effect from all other effects. Also, there should be an adequate washout period to avoid carryover effects. Usually a 2×2 COD is used for BE studies. However, other designs such as repeated COD, parallel design, etc. could be used depending on the study objective. Generally a single dose study is preferred with steady-state studies used at certain times only.

A randomised, single-dose, two-period, two sequence COD was used for this study. The T and R drugs were administered under fasting conditions (subjects are not allowed to eat 24 hours before drug administration) to the subjects with a sufficient washout period separating the two treatments. The study was done at two separate periods. In the first period of the first sequence every subject was administered a single dose (10 mg of drug) of the R drug and samples of blood were drawn at times 0, 0.5, 1, 1.5, 2, 2.5, 3, 4, 5, 6, 8, 12, 24, 30, 48 and 72 hours. The choice of the sampling times are influenced by the half-life of the drug and other pharmacokinetic parameters of the drug such as the absorption and elimination rates. In the second period each of the same subjects received another single dose (10 mg of drug) of the T drug and a new series of blood samples were drawn. Similarly, in the first period of the second sequence every subject in the second group received the T formulation first and in the second period, they received the R formulation second. In each of the periods for both sequences, samples of blood are drawn at the same times after the administration of the drug (Chow and Liu, 2008).

3.3   Research Data

Blood samples obtained through a 2×2 crossover design from pigs were used as the data for this dissertation. The 16 pigs were allocated to two different sequences. The pigs in sequence 1 were dosed with the T drug first and then with the R drug. The drugs were administered intramuscularly. The pigs in the second sequence were dosed with the R drug first and then with the T drug after the washout period. The data were obtained from Onderstepoort

Veterinary School. The study at Onderstepoort was executed to obtain BE between the T and R drugs which were both antibiotics for marketing purposes.

The dose of the drugs for the pig data is 15 mg/kg. The raw data comprise of the variables listed in Table 3.1.

Table 3.1: A list of the variables in the raw data.

Variable	Description
Time	Time at which samples are collected (hours).
Sequence	The sequence to which each subject belongs.
Period	The period to which a subject belongs.
Concentration	A measure of the concentration at a particular time.
Treatment	An R and T formulation.

Table 3.2 summarises the data used in this dissertation.

Table 3.2: Data from a Two Period COD with T and R formulations.

Subject	Sequence	AUC Test	AUC Ref	AUC T:R	C_{max} Test	C_{max} Ref	C_{max} T:R
1	RT	53.378	62.4935	0.85	4.68	3.3	1.42
2	RT	63.4214	60.244	1.05	2.68	3.16	0.85
3	RT	78.4096	65.0565	1.21	3.32	2.08	1.60
4	RT	56.0545	50.333	1.11	3.24	3.67	0.88
5	RT	38.5985	49.2165	0.78	1.82	2.06	0.88
6	RT	78.99	60.549	1.30	2.99	1.89	1.58
7	RT	71.838	57.7215	1.24	2.15	3.69	0.58
8	RT	54.3085	49.312	1.10	3.88	2.86	1.36
9	TR	56.264	55.291	1.02	5.8	1.95	2.97
10	TR	61.1015	73.1955	0.83	3.02	3.95	0.76
11	TR	59.2165	72.3501	0.82	2.12	3.26	0.65
12	TR	60.798	66.7527	0.91	3.1	10.2	0.30
13	TR	68.9234	85.6995	0.80	9.72	4.39	2.21
14	TR	67.7175	77.544	0.87	9.78	3.32	2.95
15	TR	66.153	59.1085	1.12	2.51	1.86	1.35
16	TR	58.9715	67.0836	0.88	2.37	3.69	0.64
R=Reference, T=Test							

The AUC and C_{max} data were estimated using a non-compartmental PK method and analysed using Statistical Analysis System (SAS) version 9.4, R software version 3.2.2, matrix laboratory (MATLAB) version 8 and Monolix version 2016R1. From the data, some of the observations (concentrations) were missing but were not imputed since they were only few and not likely to impact the results significantly. There was no demographic information (age, weight, height) available for the subjects in the data. There was an equal number of subjects (8 subjects per sequence) in each of the sequences which

makes it a balanced design.

3.4   Sample size

The number of subjects to be used in a BE study to establish bioequivalence between the T and R drugs within meaningful limits is determined by various factors (Chow and Liu, 2008):

- the level of significance desired;

- the required power; and

- the expected deviation from the R drug which is compatible with bioequivalence.

The analytical and clinical standards required may also determine the total number of subjects to be used. However, the minimum number of subjects used for BE is 12. An adequate number of subjects are required to be used to accommodate for any dropouts since replacing subjects during the study could complicate the statistical model and analysis. In addition, the number of subjects should also be adequate to ensure that statistically meaningful analyses and results are obtained. Table 3.3 shows the minimum sample size requirements for selected countries.

Table 3.3: The minimum samples required in some selected countries.

Regulatory Authority	Minimum Number	Sample size specification
South Africa	12 or more subjects for immediate release and 20 for modified release oral dosage forms.	The number of subjects should be justified on basis of providing 80% power.
USA	12.	The number of subjects needs to provide sufficient power to conclude bioequivalence.
Europe	12 or more.	The number of subjects should depend on appropriate sample size calculations.

3.4.1 Subject Selection

In BE studies, subjects must be chosen with the aim to minimise variation and also to allow for easy identification of the differences between the T and R formulations (Chow and Liu, 2008). The study, however, is usually performed using healthy subjects. Subjects which are selected for bioequivalence studies:

- are healthy subjects of either sex in the range of 18 to 55 years;

- are subjects with no history of drug or alcohol abuse;

- are of normal weight (weight corresponds to height as determined by BMI); and

- are subjects that have gone through thorough medical screening before selection (FDA, 2003).

However, the subject selection normally differs from one country to the next as indicated in Table 3.4.

Table 3.4: Subject selection criteria in selected countries.

Regulatory Authority	Age	Gender	BMI
South Africa	18 to 55 years.	Male or female.	Recommended BMI or usually within 15% of the body mass.
USA	18 years or older.	Male or female.	Not specified.
Europe	18 years or older.	Male or female.	18.5 to 30kg/m^2.

When animals are used in BE studies they must be healthy and chosen from a homogeneous population. Thus, they need to have the same weight, be of the same breed, same age and same sex. However, when it is difficult to get a homogeneous group of animals, it is acceptable to use animals from different populations provided age, sex and weight are controlled (EMEA, 2001).

3.4.2 Standardisation of the study

The subjects involved in the BE studies should preferably fast at least the night before the administration of the drugs or formulations. The time of the day where the subjects ingest the drug should be stated. Subjects should not be permitted to take any other drug before and during the study period.

3.5 Research Design

3.5.1 2 x 2 Crossover Design

A COD is a redesigned, randomized block in which every block receives different formulations of the drug at different periods (Chow and Liu, 2008). Crossover trials are designed in a way such that subjects receive a sequence of

treatments in a specified order for a pre-specified length of time with a washout period between each treatment where no treatments are dispensed (Wang and Bakhai, 2006). A crossover study should be designed such that the formulation effect could be separated from other types of effects (Hauschke *et al.,* 2007). In a COD two or more treatments are compared where a subject switches to another treatment upon completion of one. This is one advantage a COD trial has over the parallel design in that subjects become their own controls thereby requiring fewer subjects to evaluate the effect of different therapies. Subjects are assumed to have stable conditions which should not vary between the first and second dosing periods. This assumption must be satisfied for a COD to be implemented successfully. In addition to this, a COD is favourable for BE studies because:

- every subject serves as its own control thereby enabling each formulation to be compared within each subject and evaluating the within subject formulation;

- inter-subject variability is accounted for; and

- a good and unbiased estimate of the differences between the two formulations is obtained when the subjects are properly randomised.

A COD has the advantage of comparing the individual effects of the treatment and not the sequences. It allows for identifying the individual reactions of the various formulations as each subject receives each of the formulations (Senn, 2008). Depending upon the different number of treatments, the aim of the study and the number of sequences, there could be many sets of possible sequences which could be utilised in the design.

However, the two-period, two-treatment COD, where every subject is administered either the T or the R drug formulation first and the alternative as the second treatment, remains the simplest design. It is not compulsory that the number of periods should be the same as the number of formulations. The

treatment is randomly administered to the subjects. Most often, half the subjects receive one treatment while the other half receives the second treatment at the same time. For example, in a 2 × 2 COD, each subject is assigned arbitrarily to either treatment sequence RT or TR at different dosing periods. Subjects in the COD are administered the R and T formulations in both periods 1 and 2, respectively, in the first sequence whereas the second group of subjects also receive the T and R formulations in periods 1 and 2, respectively, in the second sequence. That is, subjects in the RT sequence receive R for the first dose and T as the second dose separated by a washout period, while subjects in the TR sequence receive T for the first dose and R as the second dose.

The main advantage of a COD trial is the comparison of treatments within subjects (Jones and Kenward, 2014). In this study, for example, each subject provides two measurements in R and T, separately. A repeated measures design has a lot of advantages and a potential disadvantage as well, but it is not considered in this dissertation. The disadvantage of a repeated design (carryover or residual effect) is the chance of a treatment effect in a period to be still present during the start of the next period. This disadvantage, however, can be reduced when the appropriate design and analysis are used (Jones and Kenward, 2014). To eliminate this carryover or residual effect a washout period of sufficient length is usually allowed.

Differences between treatments in one period could also be different from a later period because of treatment-by-period relations. However, in a well planned design, the chances of treatment-by-period effect will be minimal. A suitable design, which allows any such interaction to be easily detected and determine whether it was as a result of carryover or not, could be used as well (Jones and Kenward, 2014).

A 2×2 crossover design is shown in Figure 3.1 below.
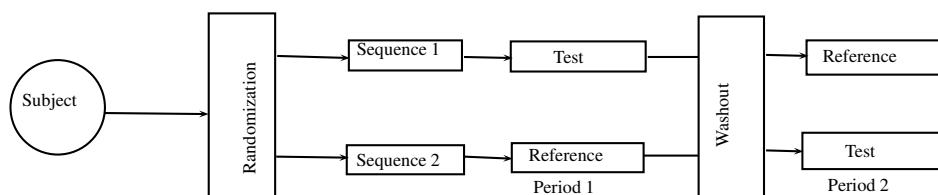
Figure 3.1: A standard two-sequence, two-period crossover design.

Generally, a clinical trial with a sequences of treatments delivered over b different periods is referred to as an $a \times b$ COD. In a COD, treatment differences are based on within-subject comparison which results in a low variability within the subjects as compared to a parallel design where treatment differences are based on between-subject variability resulting in high variability between subjects.

COD, however, poses a lot of inconvenience to subjects such as pain if it is intravenous as they have to take multiple treatments which involves a longer period of observation. This could result in high drop out rates resulting in an unequal number of subjects completing the trial. One of the challenges of a COD is the carryover effect (Jones and Kenward, 2014). Carryover effect is the tenacity of a treatment which is given in one treatment stage of the study to the next treatment stage or the effect of a drug that remains after the dosing period is over (Senn, 2008; Chow and Liu, 2008). The carryover effects (residual effect) and washout should be established in a COD as the existence of a carryover effect normally affects the inferences that are made about the BA between the drug formulations (Chow and Liu, 2008). Carryover effects may lead to biased estimation of the PK parameters.

The washout period should be long enough for the drug formulations to be completely eliminated from the subjects' system to avoid carryover effects. Washout periods normally depend on the type of drug being used in the BE studies. A first-order carryover effect lasts for one treatment period. The

period effect is another potential drawback of the COD. The effect of either the T or R treatment could be affected if it is administered either in the first or second period. The period effect could be minimised by assigning an equal number of subjects to each sequence and then adjusted statistically (Jones and Kenward, 2014).

3.5.2 Statistical Model

The two sequences, two periods COD is mostly recommended for bioequivalence assessment using PK studies. A 2×2 crossover model, as presented by Chow and Liu (2008), is given as follows:

$$Y_{ijk} = \mu + S_{ik} + P_j + F_{j,k} + C_{j-1,k} + e_{ijk}, \qquad (3.1)$$

where μ represents the overall mean; S_{ik} represents the random effect for each subject i in the kth sequence, where i=1,2...n_k, k=1,2; P_j represents the fixed effect of the jth period with j=1,...,p and $\Sigma P_j = 0$; $F_{j,k}$ is the fixed effect of the drug formulations in the kth sequence dispensed during the jth period while $\Sigma F_{j,k} = 0$; $C_{j-1,k}$ is the fixed carryover effect of the formulations in the kth sequence dispensed at the $(j-1)$th period where $C_{o,k} = 0$ and $\Sigma C_{j-1,k} = 0$; e_{ijk} is the within-subject random error; and Y_{ijk} is the PK parameter of the ith subject for each sequence k at each period j.

The assumption is that S_{ik} should be independently and identically distributed with a mean of zero and variance σ_s^2, and e_{ijk} should be independently distributed with a mean of zero and variance σ_t^2, where $t = 1, 2, ..., l$ (the number of drug formulations). Carryover effects, period effects and direct drug effects are usually the preliminary tests that are performed before assessing bioequivalence between the T and R formulations.

All the methods which are used in this study for BE assessment are derived using the model in Equation 3.2,

$$Y_{ijk} = \mu + S_{ik} + F_{j,k} + P_j + e_{ijk}, \qquad (3.2)$$

where Y_{ijk}, μ, S_{ik}, $F_{j,k}$, P_j and e_{ijk} are defined in Equation 3.1.

3.5.3   Washout and Carryover Effect

Despite the advantage of each subject serving as its own control, COD still has a disadvantage. This potential problem arises when the first treatment effect from the first time period is still existing in the next period and therefore misrepresents the effect of the second treatment. According to Senn (2001) a carryover effect arises whenever the drug treatment administered in one period continues to affect subjects in the second period thereby biasing the amount of drug in the blood sample of a subject. One potential cause of a carryover effect is the short time between periods before administering the second drug formulation. The short time between the periods is referred to as a washout period. Carryover effect could also be as a result of some drugs having longer half-lives or some subjects having very weak metabolism rates. Washout and carryover effects (residual effects) need to be addressed in a COD because the carryover effect affects statistical inference of BA between formulations (Chow and Liu, 2008).

Whenever the carryover effect is present, the estimate ($\mu_T - \mu_R$) would be biased as the effect of the drug which is dispensed in the first period is still present. A washout period is the resting period between drugs dispensed in the first period and second period during which the impact of a drug formulation given at one period does not extend to the next treatment period (Chow and Liu, 2008).

The duration of the washout period is determined depending on the type of drug and half-life of the drug. For drugs with a long half-life, their effect may

still be present even after washout is introduced at the termination of a dosing period. If this happens it becomes important to differentiate the direct drug effect from the carryover effect. A direct drug effect can be interpreted as the effect of the drug being present in the period in which it was dispensed while carryover effect is the drug effect after the termination of a dosing period (Chow and Liu, 2008).

In the absence of carryover effects due to adequate washout period, the general crossover model given in Equation 3.1 reduces to the model in Equation 3.2.

3.5.4 Period Effect

It is preferred that the subjects' condition and ability to respond to the formulations remain unchanged between the two dosing periods. This is not always the case as some subjects improve on average or deteriorate before taking the second treatment. This can cause a problem which affects COD, generally referred to as the period effect (treatment by period interaction). The effect of R and T can be influenced depending on the period they were administered. This could be as a result of physiological and environmental changes which occur in the subjects. A significant period effect could be as a result of timing and amount of physical activity and/or the hotness or coldness of the water used at both periods (Lockyer, Al-Dgither, Al-Gaai, Yousuf and Hammami, 2005). However, a significant period effect does not influence the BE assessment.

3.6 Outliers

One of the challenges encountered in BA and BE studies is that the data set sometimes contains either extremely small or extremely large observations which are referred to as outliers. An outlier is a data point that does not fit the model correctly. These extremely small or extremely large observations tend to influence any decisions reached about the bioequivalence of the two drugs. There are four types of outliers as identified by (Chow and Liu, 2008):

- unexpected observations which appear in the CTC;

- very small or large observations in a given formulation type;

- subjects unusually having either extremely low or extremely high BA with respect to R form; and/or

- unusual subjects having extreme BA in both formulations.

Chow and Tse (1991) argued that the decision on bioequivalence is heavily influenced by the presence of any outliers. When BA studies are conducted with the inclusion of subjects which are potential outliers, Westlake (1981) and Schuirmann (1987) stated that this may lead to bioinequivalence whereas the formulations are indeed bioequivalent. It is therefore very important to identify any possible outliers in the data and conduct the bioequivalence studies with and without these outliers.

By doing so it makes it easy to conclude if the possible outlier does have any effect. Rodda (1986) suggested that outliers (unexpected observations) should be included in the analysis as they have little effect on the calculation of the AUC and, consequently, have little effect on BA comparison. It only suggests that the subjects respond to the formulations differently either due to external or internal factors. The FDA also believes that outliers should not be excluded from the analysis (FDA, 2001).

3.6.1 Outlier Detection

The study of the detection of potential outliers in a data set has received much attention in past decades (Dixon, 1953; Cook, 1977; Draper and John, 1981). Various methods have been developed to handle outliers in linear regression problems. However, the data used in BA studies can not be accommodated in a linear regression framework. The inclusion or deletion of potential outliers in the BA analysis affects the results obtained and as such need to be considered. A number of possible techniques have been proposed for the detection of outliers in BA and BE studies. Cook and Weisberg (1982) proposed the Likelihood Distance (LD) approach for identifying outliers while Chow and Tse

(1990) developed an approach known as the Estimates Distance (ED) method. Another approach, the Hotelling T^2 statistic (HT), was introduced by Liu and Weng (1991) to detect outliers.

3.6.1.1 Likelihood Distance and Estimates Distance

The LD and the ED procedures which were proposed to detect an outlier assume no period effect and no formulation effect. In the absence of formulation and period effects, the model in Equation 3.1 is reduced to Equation 3.3:

$$Y_{ij} = \mu + S_i + e_{ij} \quad j = 1, ..., n; \; i = 1, ..., k, \tag{3.3}$$

where n is the number of formulations, k is the number of subjects and Y_{ij} expresses the measurement value for the ith subject, jth formulation. The parameters which are of interest include μ, σ_e^2 and σ_s^2. Let $\hat{\boldsymbol{\theta}} = (\theta_1, \theta_2, \theta_3)^T$ where $\theta_1 = \mu$, $\theta_2 = \sigma_e^2$ and $\theta_3 = \sigma_e^2 + n\sigma_a^2$. The log-likelihood function L(θ) for Equation 3.3 is given as:

$$L(\theta) = \frac{-kn}{2}log2\pi - \frac{k}{2}log(\theta_2\theta_3^{n-1}) - \frac{1}{2\theta_3}\sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \theta_1)^2 - \frac{n}{2}\left(\frac{1}{\theta_2} - \frac{1}{\theta_3}\right)\sum_{i=1}^{k}(\bar{Y}_i - \theta_1)^2. \tag{3.4}$$

The maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}$, of θ is obtained through maximisation of $L(\theta)$ in Equation 3.4 with respect to θ given that $\theta_3 \geq \theta_2$. The MLE $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)^T$ of θ is then,

$$\hat{\theta}_1 = \bar{Y} = \frac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n}Y_{ij},$$

$$\hat{\theta}_2 = m_1,$$

$$\hat{\theta}_3 = \frac{(k-1)m_2}{k},$$

where

$$m_1 = \frac{1}{k(n-1)}\sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_i)^2,$$

74

and

$$m_2 = \frac{n}{k-1} \sum_{i=1}^{k} (\bar{Y}_i - \bar{Y})^2.$$

However, if $\hat{\theta}_3 < \hat{\theta}_2$, the MLE of θ_2 and θ_3 becomes

$$\hat{\theta}_2 = \hat{\theta}_3 = \frac{1}{nk} \sum_{i=1}^{k} \sum_{j=1}^{n} (Y_{ij} - \bar{Y})^2,$$

which is obtained by maximising $L(\theta)$ under the condition that $\theta_2 = \theta_3$.

Since the log-likelihood function gives a summary of the information concerning θ for a given data set, it becomes possible to compute the impact of the ith case based on the difference of the log-likelihood between $\hat{\theta}$ and $\hat{\theta}_i$, where $\hat{\theta}_i$ represents the MLE of θ when the ith subject is deleted. The difference is given as

$$L(\hat{\theta}) - L(\hat{\theta}_i), \tag{3.5}$$

which measures the impact of the ith subject. If the contrast in Equation 3.5 is large (could influence bioequivalence decision), the ith subject is considered an outlier.

The LD statistic, which is based on the log-likelihood distance, is given by

$$LD_i(\hat{\theta}) = 2[L(\hat{\theta}) - L(\hat{\theta}_i)],$$

where $\hat{\theta}_i$ is the MLE of θ when the ith subject is deleted. $LD_i(\hat{\theta})$ can be shown to be asymptotically chi-square distributed with 3 degrees of freedom (df) as k (number of subjects) approaches infinity. An ith subject is therefore considered an outlier if

$$LD_i(\hat{\theta}) > \chi_3^2(\alpha),$$

where $\chi_3^2(\alpha)$ is the upper α percentile of a central chi-square distribution with 3 df.

Alternatively, the impact of an individual subject could be studied by deleting that subject and taking note of the changes (Chow and Tse, 1990). The comparison is then based on the difference $\hat{\theta}_i - \hat{\theta}$. A squared distance from $\hat{\theta}_i$ to $\hat{\theta}$, which is referred to as Cook's distance, and is similar to that of linear regression, was derived by Cook and Weisberg (1982).

The estimates distance can then be defined as

$$ED_i(\hat{\theta}) = k^2(\hat{\theta}_i - \hat{\theta})\hat{\Sigma}^{-1}(\hat{\theta}_i - \hat{\theta}),$$

where $\hat{\Sigma}$ is the estimate of

$$\sum = \begin{bmatrix} \dfrac{\theta_3}{n} & 0 & 0 \\ 0 & \dfrac{2\theta_2^2}{n-1} & 0 \\ 0 & 0 & 2\theta_3^2 \end{bmatrix}, \tag{3.6}$$

obtained by replacing θ by its MLE $\hat{\theta}$. $ED_i(\hat{\theta})$ can be shown to have a chi-square distribution with 3 df. Therefore, an ith subject is regarded as an outlier if

$$ED_i(\hat{\alpha}) > \chi_3^2(\alpha),$$

where $\chi_3^2(\alpha)$ is the α percentage point of a χ_3^2 distribution.

3.7   Interval Hypothesis Approach

In reality, no two drug formulations can ever have the same bioavailability profiles as their profiles would differ by a clinically meaningful limit (Chow and Liu, 2008). Even though the two drug formulations may not have exactly the same profiles, they could be considered as equivalent if they differ by a clinically accepted limit. Due to this, another approach for assessing average bioequivalence, the interval hypothesis method, was introduced by Schuirmann (1987). The hypothesis for average bioequivalence is given as

$$H_0 : \mu_T - \mu_R \leq \theta_L \quad or \quad H_0 : \mu_T - \mu_R \geq \theta_U,$$
$$H_1 : \theta_L < \mu_T - \mu_R < \theta_U, \tag{3.7}$$

where μ_T and μ_R are the average bioavailability of the test product and reference product, respectively, θ_U and θ_L are clinically meaningful limits which are both given as 20% of the unknown reference mean (μ_R). The null hypothesis states that μ_T and μ_R are not equivalent while the alternative hypothesis states that they are equivalent. Bioequivalence is concluded if both null hypotheses are rejected. However, when the natural logarithm of the PK responses is used, the hypothesis in Equation 3.7 becomes:

$$H_0 : \mu_T/\mu_R \leq \delta_L \quad or \quad H_0 : \mu_T/\mu_R \geq \delta_U,$$
$$H_1 : \delta_L < \mu_T/\mu_R < \delta_U, \tag{3.8}$$

where $\delta_L = exp(\theta_L)$ and $\delta_U = exp(\theta_U)$.

Given that $\theta_L = -0.2\mu_R$ and $\theta_U = 0.2\mu_R$, the interval hypotheses can be rephrased as

$$H_0 : \mu_T - \mu_R \leq -0.2\mu_R \quad or \quad H_0 : \mu_T - \mu_R \geq 0.20\mu_R,$$
$$H_1 : -0.2\mu_R < \mu_T - \mu_R < 0.20\mu_R, \tag{3.9}$$

which, if $\mu_R > 0$, may be changed to

$$H_0 : \mu_T/\mu_R \leq 0.80 \quad or \quad H_0 : \mu_T/\mu_R \geq 1.20,$$
$$H_1 : 0.80 < \mu_T/\mu_R < 1.20. \tag{3.10}$$

The statistical assumptions of normality and homogeneity of variance are deemed to be satisfied for the logarithmically transformed bioavailability variables (Hauck and Anderson, 1984). If the statistical assumptions are true on the logarithmic scale, the interval hypotheses can be changed to

$$H_0 : \eta_T - \eta_R \leq log(0.80) \quad or \quad H_0 : \eta_T - \eta_R \geq log(1.20),$$
$$H_1 : log(0.80) < \eta_T - \eta_R < log(1.20), \tag{3.11}$$

where η_T and η_R are the true test and reference means, respectively, of the logarithmically transformed parameters.

Equation 3.7 can be decomposed into:

$$H_{01} : \mu_T - \mu_R \leq \theta_L \qquad versus \qquad H_{a1} : \mu_T - \mu_R > \theta_L,$$
$$H_{02} : \mu_T - \mu_R \geq \theta_U \qquad versus \qquad H_{a2} : \mu_T - \mu_R < \theta_U. \tag{3.12}$$

The initial set of hypotheses in Equation 3.12 are used to determine if average bioavalibality of the T drug is not too low while the second batch tests if the average bioavailability is not too high (Chow and Liu, 2008). Whenever the average BA of the T drug formulation is too low, there is an issue of efficacy and an issue of safety when it is too high. Rejecting H_0 in Equation 3.7 is similar to rejecting H_{01} and H_{02} in Equation 3.12 which eventually leads to an average bioequivalence conclusion.

Thus, when $\theta_L < \mu_T - \mu_R$ and $\mu_T - \mu_R < \theta_U$, it can be concluded that

$$\theta_L < \mu_T - \mu_R < \theta_U,$$

μ_T and μ_R are considered to be bioequivalent.

3.7.1 Two One-Sided Test (TOST)

The logic of hypothesis testing requires that the hypothesis to be demonstrated be the alternative hypothesis. This, however, implies that the equivalence hypothesis should be the alternative and not the null hypothesis as demonstrated in Equation 3.7. Schuirmann (1987) introduced the Two One-Sided Test (TOST) procedure which consists of decomposing the interval hypotheses in Equation 3.7 into two sets of one-sided hypotheses,

$$
\begin{aligned}
H_{01} &: \mu_T - \mu_R \leq \theta_L & versus & & H_{a1} &: \mu_T - \mu_R > \theta_L, \\
H_{02} &: \mu_T - \mu_R \geq \theta_U & versus & & H_{a2} &: \mu_T - \mu_R < \theta_U.
\end{aligned}
\tag{3.13}
$$

The TOST procedure consists of rejecting the null hypothesis in the interval hypothesis (Equation 3.13) and concluding equivalence of μ_T and μ_R, if and only if both null hypotheses (H_{01} and H_{02}) are rejected at a specified α level of significance. Under normality assumptions, the two sets of one-sided hypotheses in Equation 3.13 can be tested with ordinary one-sided t tests. Thus, BE is concluded if

$$T_L = \left(\frac{(\bar{Y}_T - \bar{Y}_R) - \theta_L}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right) > t(\alpha, n_1 + n_2 - 2),$$

$$\text{(3.14)}$$

$$T_U = \left(\frac{(\bar{Y}_T - \bar{Y}_R) - \theta_U}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right) < -t(\alpha, n_1 + n_2 - 2),$$

where $\bar{Y}_T$ and $\bar{Y}_R$ are the observed average bioavailabilities of the T and R formulations, respectively, θ_L and θ_U are lower and upper bioequivalent limits, $\hat{\sigma}_d$ is the pooled standard deviation of period difference from the two sequences, n_1 and n_2 are the number of samples in each sequence and α is the level of significance. The TOST procedure turns out to be operationally identical to the procedure of declaring bioequivalence only if the $(1 - 2\alpha) \times 100\%$ CI for $\mu_T - \mu_R$ is contained in the interval $[\theta_L, \theta_U]$, where $(1 - 2\alpha) \times 100\%$ CI is the 90% classical confidence interval.

However, regulatory agencies such as the FDA were not necessarily concerned with the power of bioequivalence but rather with the confidence interval with which bioequivalence is concluded (Patterson, 2003). According to FDA (1992), the significance level under the Schiurman's TOST procedure is set at 5% per test. The assumptions for bioequivalence assessment using the TOST procedure are as follows:

- Samples should be randomised;

- Homogeneity of variances;

- Addition of the statistical model; and

- Normality and independence of residuals.

The FDA demands that subjects used for the bioequivalence studies must be allocated at random to a sequence and the model will be fitted to a log transformed AUC and C_{max}. The within- and between-subjects variances are

79

assumed to be homogeneous for both formulations while the C_{max} and AUC data are presumed to be log-normally distributed. The goodness-of-fit of the log-transformed data is assessed using the residuals versus predicted values as well as normality probability plots (Jones and Kenward, 2014).

3.7.2 Analysis of Variance (ANOVA)

ANOVA has been used as an assessment of bioequivalence in the past, according to Pabst and Jaeger (1990), to test a simple hypothesis of the means of the characteristics of the CTC being equal or not (Westlake, 1981). The null hypothesis states that the T drug has the same consequence as the R drug. Rejecting the null hypothesis implies that the T and the R have different effects whenever they are taken. The hypothesis is given as follows:

$$H_0 : \mu_i - \mu = 0 \quad for \quad all \quad i = 1, 2, ..k$$
$$H_1 : at \quad least \quad two \quad \mu_i - \mu \quad are \quad different,$$

(3.15)

where μ_i represents the population mean for level i. Quite a handful of researchers have criticised the use of the ANOVA as a method of bioequivalence assessment (Hauck and Anderson, 1984). However, the ANOVA is needed, as stated by Rani and Pargal (2004), to estimate the least squares means and the error variance, σ^2. The use of ANOVA enables the researcher to comprehend the variation within the data through separating the total sum of squares (SS) into random errors and fixed effects (Chow and Liu, 2008). The model used in the ANOVA, given by Chow and Liu (2008), is given in Equation 3.16

$$
\begin{aligned}
SS_{Total} &= \sum_{k=1}^{2} \sum_{j=1}^{2} \sum_{i=1}^{n_k} (Y_{ijk} - \bar{Y}_{...})^2 \\
&= \sum_{k=1}^{2} \sum_{j=1}^{2} \sum_{i=1}^{n_k} (Y_{ijk} - \bar{Y}_{i.k} + \bar{Y}_{i.k} - \bar{Y}_{...})^2 \\
&= \sum_{k=1}^{2} \sum_{j=1}^{2} \sum_{i=1}^{n_k} (Y_{ijk} - \bar{Y}_{i.k})^2 + 2 \sum_{k=1}^{2} \sum_{i=1}^{n_k} (Y_{i.k} - \bar{Y}_{...})^2 \\
&= SS_{within} + SS_{between},
\end{aligned}
$$

(3.16)

where

$$\bar{Y}_{i.k} = \frac{1}{2} \sum_{j=1}^{2} Y_{ijk},$$

where $\bar{Y}_{i.k}$ is the average of Y_{ijk} for each sequence, $\bar{Y}_{...}$ is the grand mean, SS_{within} denotes the sum of squares for within subjects and $SS_{between}$ is the sum of squares between the subjects. The number of subjects in sequences 1 and 2 are denoted as n_1 and n_2, respectively. The total sum of squares (SS_{total}) has $2(n_1 + n_2) - 1$ df, the SS_{within} has $n_1 + n_2$ df and the $SS_{between}$ has $n_1 + n_2 - 1$ df. The $SS_{between}$ is sub-divided into two parts with one being the carryover effect and the other the inter-subject error. Thus,

$$SS_{between} = SS_{carry} + SS_{inter}, \tag{3.17}$$

with SS_{carry} being the carryover effects with 1 df and SS_{inter} the inter-subject error with $n_1 + n_2 - 2$ df. SS_{carry} and SS_{inter} are defined below,

$$\boldsymbol{SS}_{carry} = \frac{2n_1 n_2}{n_1 + n_2} \{(\bar{Y}_{.12} + \bar{Y}_{.22}) - (\bar{Y}_{.11} + \bar{Y}_{.21})\}^2, \quad and$$

$$\boldsymbol{SS}_{inter} = \sum_{k=1}^{2} \sum_{i=1}^{n_k} \frac{Y_{i.k}^2}{2} - \sum_{k=1}^{2} \frac{Y_{..k}^2}{2n_k},$$

The mean squares (MS) is derived by dividing the sum of squares by its df. To test the carryover effect hypothesis, the test statistic is given by

$$F_c = \frac{MS_{carry}}{MS_{inter}} \sim F(1, n_1 + n_2 - 2).$$

If $F_c < F(\alpha, 1, n_1 + n_2 - 2)$, the null hypothesis is rejected with $F(\alpha, 1, n_1 + n_2 - 2)$ being the upper αth percentile for the F distribution with 1 and $n_1 + n_2 - 2$ df (Chow and Liu, 2008).

However, the test statistic F_c is similar to $T_c > t(\alpha/2, n_1 + n_2 - 2)$ since $F_c = T_c^2$. In the same way, SS_{within} can be written as:

$$SS_{within} = SS_{drug} + SS_{period} + SS_{intra},$$

where

$$SS_{drug} = \frac{2n_1 n_2}{n_1 + n_2} \{ \frac{1}{2} [(\bar{Y}_{.21} - \bar{Y}_{.11}) - (\bar{Y}_{.22} - \bar{Y}_{.12})] \}^2,$$

$$SS_{period} = \frac{2n_1 n_2}{n_1 + n_2} \{ \frac{1}{2} [(\bar{Y}_{.21} - \bar{Y}_{.11}) - (\bar{Y}_{.12} - \bar{Y}_{.22})] \}^2,$$

$$SS_{intra} = \sum_{k=1}^{2} \sum_{j=1}^{2} \sum_{i=1}^{n_k} Y_{ijk}^2 - \sum_{k=1}^{2} \sum_{i=1}^{n_k} \frac{Y_{i.k}^2}{2} - \sum_{k=1}^{2} \sum_{j=1}^{2} \frac{Y_{.jk}^2}{n_k} + \sum_{k=1}^{2} \frac{Y_{..k}^2}{2n_k},$$

with 1 df for each of SS_{drug} and SS_{period} and $n_1 + n_2 - 2$ df for SS_{intra}.

Whenever the null hypothesis for the carryover effect is not rejected, thus $C_R = C_T$, the null hypothesis for the drug effect testing for no drug effect can be tested using the statistic

$$F_d = \frac{MS_{drug}}{MS_{intra}} \sim F(1, n_1 + n_2 - 2).$$

The null hypothesis is rejected if $F_d > F(\alpha, 1, n_1 + n_2 - 2)$. The test statistic F_d is similar to

$$T_d = \frac{\hat{F}}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

since $F_d = T_d^2$.

The test statistic below is considered for testing the null hypothesis in a period effect:

$$F_p = \frac{MS_{period}}{MS_{intra}} \sim F(1, n_1 + n_2 - 2).$$

The null hypothesis is then rejected if $F_p > F(\alpha, 1, n_1 + n_2 - 2)$. It could be shown that

$$F_p = T_0 = \frac{\hat{P}}{\hat{\sigma}_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

An ANOVA table for the 2×2 COD is provided in Table 3.5.

Table 3.5: Analysis of variance table for the 2×2 cross-over design.

Source of variation	DF	Sum of Squares(SS)	$MS = SS/df$	F
Inter-Subjects				
Carryover	1	SS_{carry}	SS_{carry}	$F_C = MS_{carry}/MS_{inter}$
Residual	$n_1 + n_2 - 2$	SS_{inter}	$SS_{inter}/n_1 + n_2 - 2$	$F_V = MS_{inter}/MS_{intra}$
Intra-subjects				
Direct Drug	1	SS_{drug}	SS_{drug}	$F_d = MS_{drug}/MS_{intra}$
Period	1	SS_{period}	SS_{period}	$F_P = MS_{period}/MS_{intra}$
Residual	$n_1 + n_2 - 2$	SS_{intra}	$SS_{intra}/n_1 + n_2 - 2$	
Total	$2(n_1 + n_2) - 1$	SS_{total}		

However, to test for inter-subject variability (σ_s^2), the following hypothesis is considered:

$$H_0 : \sigma_s^2 = 0,$$
$$H_0 : \sigma_s^2 > 0. \tag{3.18}$$

The test statistic for testing the inter-subject variability is

$$F_v = \frac{MS_{inter}}{MS_{intra}} \sim F(1, n_1 + n_2 - 2, n_1 + n_2 - 2).$$

If $F_v > F(\alpha, n_1 + n_2 - 2, n_1 + n_2 - 2)$, the null hypothesis of no inter-subject variation is rejected.

3.7.3 Wilcoxon-Mann-Whitney Test

A distribution-free rank sum test can sometimes be used on a TOST, according to Hauschke *et al.* (2007), with a 2×2 COD (Chow and Liu, 2008). This distribution-free rank sum test is known as a Wilcoxon-Mann-Whitney TOST (Chow and Liu, 2008). This is a non-parametric approach to the parametric approach by Schuirmann (1987). When $\mu_T - \mu_R$ in the hypothesis stated in Equation 3.13 is replaced by θ, the two sets of hypotheses are re-written as:

$$H_{01} : \theta_L^* \leq 0 \qquad versus \qquad H_{a1} : \theta_L^* > 0$$
$$H_{02} : \theta_U^* \geq 0 \qquad versus \qquad H_{a2} : \theta_U^* < 0, \tag{3.19}$$

where $\theta_L^* = \theta - \theta_L$ and $\theta_U^* = \theta - \theta_U$.

Given that $i = 1, 2, ..., n_k$ and $k = 1, 2$, the approximations of θ_L^* and θ_U^* are acquired as a linear function of the differences in the period d_{ik}. Let:

$$b_{hik} = \begin{cases} d_{ik} - \theta_h, & h = U, L, \quad \text{for the subjects in sequence 1,} \\ d_{ik}, & \text{for subjects in sequence 2,} \end{cases} \tag{3.20}$$

where $d_{ik} = \frac{1}{2}(Y_{i2k} - Y_{i1k})$. In the absence of any carryover effect, the expectation of b_{hik} is given as:

$$E(b_{hik}) = \begin{cases} \frac{1}{2}[(P_2 - P_1) + (\theta - 2\theta_h)], & \text{for k=1,} \\ \frac{1}{2}[(P_2 - P_1) - \theta], & \text{for k=2,} \end{cases} \tag{3.21}$$

84

with $h = L, U$, $i = 1, 2, ..., n_k$ and $k = 1, 2$. The variance is given by $V(b_{hik}) = V(d_{ik}) = \sigma_d^2 = \frac{\sigma_e^2}{2}$. It, however, implies that $E(b_{hi1}) - E(b_{hi2}) = (\theta - \theta_h) = \theta_h^*$. Let R_L represent the rank sum of the responses for subjects in the sequence 1,

$$R_L = \sum_{i=1}^{n1} R(b_{Li1}). \tag{3.22}$$

The test statistic for H_{01} is given by:

$$W_L = R_L - \frac{n_1(n_1 + 1)}{2}. \tag{3.23}$$

If $W_L > W(1 - \alpha)$, the null hypothesis (H_{01}) is rejected. $W(1 - \alpha)$ represents the $(1 - \alpha)^{th}$ quantile for the distribution of W_L. Similarly, for the second batch of hypotheses in Equation 3.19, H_{02} is rejected if:

$$W_U = R_U - \frac{n_1(n_1 + 1)}{2} < W(\alpha), \tag{3.24}$$

where R_U represents the sum of ranks of b_{Uik} for the subjects in the first sequence. When both sets of hypotheses H_{01} and H_{02} are rejected, thus

$$W_L > W(1 - \alpha) \quad and \quad W_U < W(\alpha), \tag{3.25}$$

the two drugs are considered to be average bioequivalent. Whenever there are no ties among the observations, the expectation and variances of W_L and W_U, under the null hypotheses H_{01} and H_{02}, gives

$$E(W_L) = E(W_U) = \frac{n_1 n_2}{2},$$
$$V(W_L) = V(W_U) = \frac{1}{12}(n_1 + n_2 + 1). \tag{3.26}$$

However, when there are ties, the W_L and W_U are calculated by assigning average ranks. The expected value and variances of W_L and W_U thus become

$$E(W_L) = E(W_U) = \frac{n_1 n_2}{2}, \tag{3.27}$$

and

$$V(W_L) = V(W_U) = \frac{1}{12}(n_1 + n_2 + 1 - Q), \tag{3.28}$$

with

$$Q = \frac{1}{(n_1 + n_2)(n_1 + n_2 - 1)} \sum_{v=1}^{q} R(r_v^3 - r_v),$$

85

where q represents the frequency of groups that are semantic while r_v denotes the size of the groups that have the same association v.

It could be concluded that $W(1 - \alpha) = n_1 n_2 - W(\alpha)$ since W_L and W_U are symmetric about their means, $n_1 n_2 / 2$. For a large total number of subjects (i.e $n_1 + n_2 > 40$) and with the ratio of n_1 to n_2 close to $1/2$, Equation 3.25 can be approximated for ABE with a large sample approximation which uses the standard normal distribution. Bioequivalence could be established if $Z_L > Z(\alpha)$ and $Z_U < -Z(\alpha)$ with $Z(\alpha)$ being the αth quantile of the standard normal distribution and

$$Z_L = \frac{W_L - E(W_L)}{\sqrt{V(W_L)}} = \frac{R_L - \left[\dfrac{n_1(n_1 + n_2 + 1)}{2}\right]}{\sqrt{\dfrac{1}{12}n_1 n_2(n_1 + n_2 + 1)}},$$

$$Z_U = \frac{W_U - E(W_U)}{\sqrt{V(W_U)}} = \frac{R_U - \left[\dfrac{n_1(n_1 + n_2 + 1)}{2}\right]}{\sqrt{\dfrac{1}{12}n_1 n_2(n_1 + n_2 + 1)}}.$$

(3.29)

The variances in Z_L and Z_U are replaced with the one given in Equation 3.28 which assumes no ties.

Statistical methods for assessing average BE between the R and T drugs are derived with the assumption that S_{ik} as well as e_{ijk} are normally distributed with means of 0 and variances σ_s^2 and σ_e^2. In practise, the challenge faced when comparing two drugs is the assumption of normality. Whenever the assumption of normality is violated for both the raw or log-transformed data, a distribution-free (non-parametric) method is applied. The non-parametric version of TOST, which is the Wilcoxon-Mann-Whitney TOST, would be used to test for a difference between T_{max} for the R and the T drugs. The statistical inferences discussed above assumed that the data follow a normal distribution. The parametric test can then be replaced by Wilcoxon rank-sum tests or the Mann-Whitney U-test. The non-parametric analysis for a 2×2 COD was described first by Koch (1972) and illustrated later by Cornell (1990). In order

to use this approach, the subjects must be randomly selected from a population of healthy subjects, the observations within a sample must be independent of one another and the two subjects must be independent of one another.

3.7.4 Confidence Interval

For some years, BE studies have been simply done using ANOVA, where a simple hypothesis of no difference in the means of the R drug and T drug was tested (Westlake, 1981). The use of the theory of classical statistical hypothesis testing of the equality of any two formulations has been criticised by many researchers, such as Westlake (1972) and Metzler (1974), as not being the appropriate statistical method. Both of them suggested that the right approach is the confidence interval (CI) method for the assessment of average BE.

Westlake (1972) further argued that the hypothesis approach was not relevant to assess bioequivalence and proposed an alternative method which is based on the confidence interval. Also, it is not sufficient to say the two drugs are interchangeable, but that an amount of assurance should be given that the average quantity of the drug eliminated from the T drug is adequately close to the quantity eliminated by the R drug. The hypothesis method, however, is not able to provide this assurance. Metzler (1974) as well stated that it is not enough to prove that $B_r \neq B_t$ (B_r is the bioavailability of the R drug and B_t is the bioavailability of the T drug) but that the difference $B_r - B_t$ should be assessed instead and conclude it is acceptably small and thus non-significant. Classical hypotheses testing are sometimes not appropriate for BE studies. The classical hypothesis testing only tests whether the two drug formulations are identical. Hypothesis testing, however, does not provide this assurance. The confidence interval approach provides this assurance which is needed to declare bioequivalence between the R and T drug.

To establish ABE between the R and T formulation, a 90% confidence interval is constructed on either the natural logarithm or logarithm to base 10 of the

AUC and C_{max}. The two drugs can then be said to be bioequivalent if a classic $(1-2\alpha) \times 100\%$ CI for the ratio of the means of the T to the R drug μ_T/μ_R or the difference between the mean of the T drug and the mean of the R drug $(\mu_T - \mu_R)$ is within the recommended limits set by the FDA and acceptable limits of other regulatory bodies (Westlake, 1981). The acceptable limits for the ratio of μ_T/μ_R, according to the FDA (2001), are [0.8,1.25] .

Metzler (1974) and Westlake (1976) stated that the classic CI method is the most suitable approach for the assessment of ABE. Kirkwood and Westlake (1981) suggested that when the $(1-2\alpha) \times 100\%$ CI for $(\mu_T - \mu_R)$ lies within the approved limits approved by the FDA and other regulatory agencies, a conclusion could be made indicating that the T drug is bioequivalent to the R drug. However, when this condition is not satisfied, the T drug cannot be said to be bioequivalent to the R drug.

The two CI approaches which will be used in this thesis are the classic CI and Westlake's Symmetric CI approaches.

3.7.4.1 Classic Confidence Interval

Given that $\bar{Y}_R$ and $\bar{Y}_T$ represent the least squares means (the average means of each formulation in both periods) of the R and T drugs, respectively, acquired from the sequence-by-period means, the classic $(1-2\alpha) \times 100\%$ CI is deduced using the t-statistics in Equation 3.30,

$$T = \frac{(\bar{Y}_T - \bar{Y}_R) - (\mu_T - \mu_R)}{\hat{\sigma}_d\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t(n_1 + n_2 - 2), \tag{3.30}$$

where the total number of subjects in sequence 1 and sequence 2 are given by n_1 and n_2, respectively, and $\hat{\sigma}_d$ denotes the pooled standard deviation of the sample according to the Student t-distribution with $(n_1 + n_2 - 2)$ df (Chow and Liu, 2008). A classic $(1-2\alpha) \times 100\%$ CI for the mean difference, $\mu_T - \mu_R$, has limits

$$L_1 = (\bar{Y}_T - \bar{Y}_R) - t(\alpha, (n_1 + n_2 - 2))\hat{\sigma}_d\sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

$$U_1 = (\bar{Y}_T - \bar{Y}_R) + t(\alpha, (n_1 + n_2 - 2))\hat{\sigma}_d\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}. \tag{3.31}$$

Equation 3.31 could be converted to a classic $(1 - 2\alpha) \times 100\%$ approximate CI for the mean ratio, μ_T/μ_R, which results in Equation 3.32,

$$L_2 = \left(\frac{L_1}{\bar{Y}_R} + 1\right) \times 100\%,$$

$$U_2 = \left(\frac{U_1}{\bar{Y}_R} + 1\right) \times 100\%. \tag{3.32}$$

Given that θ_U and θ_L are, respectively, the upper and lower equivalence limits for the difference, and δ_U and δ_L are the upper and lower equivalence limits for the ratio, bioequivalence is concluded if

$$(L_1, U_1) \in (\theta_L, \theta_U),$$

$$(L_2, U_2) \in (\delta_L, \delta_U), \tag{3.33}$$

where:

- $\theta_L = -0.2\mu_R$;

- $\theta_U = 0.2\mu_R$;

- $\delta_L = 80\%$; and

- $\delta_U = 120\%$ for the ± 20 rule.

The classic $(1 - 2\alpha) \times 100\%$ CI for the mean difference, $\mu_T - \mu_R$, is a random interval and the confidence limits which are associated with it are random variables. The classic $(1 - 2\alpha) \times 100\%$ CI for $\mu_T - \mu_R$ has a basic concept that, should the same study be performed a number of times, say n times, then $(1 - 2\alpha) \times 100\%$ times of the n randomly constructed intervals will normally cover $\mu_T - \mu_R$ range (Peter and Kjell, 2001).

However, there is no guarantee that the probability of the classic $(1 - 2\alpha) \times$ 100% CI within the limits of the equivalence limits is at least $1 - 2\alpha$ (Chow and Liu, 2008). Thus, the probability,

$$P[(L_1, U_1) \in (\theta_L, \theta_U)],$$

cannot be automatically more than or equal to $1 - 2\alpha$.

3.7.4.2 Westlake's Symmetric Confidence Interval

The classic $(1 - 2\alpha) \times 100\%$ CI for the difference in the means of the T and R drugs, $\mu_T - \mu_R$, from Equation 3.31, is symmetric around $\bar{Y}_T - \bar{Y}_R$ but not around zero (Chow and Liu, 2008). Similarly, Chow and Liu (2008) stated that, from Equation 3.32, the CI for the mean ratio of the T and R drugs, μ_T/μ_R, is symmetric around $\bar{Y}_T/\bar{Y}_R$ but not around 1. The classic CI derived from an unpaired two sample t statistic in Equation 3.30 is

$$|T| < k \ \ or \ \ -k < T < k,$$

where k is the upper αth percentile of a central t distribution with $n_1 + n_2 - 2$ degrees of freedom. The classic $(1 - 2\alpha) \times 100\%$ CI for the difference in the means of the T and R drugs, $\mu_T - \mu_R$, could be given as:

$$k_1 < T < k_2, \tag{3.34}$$

where k_1 and k_2 are selected such that the chance from k_1 to k_2, which is dependent on a central t-distribution with $n_1 + n_2 - 2$ df, is equivalent to $(1 - 2\alpha)$. That is,

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

Equation 3.34 reduces to the classic CI Equation 3.31 whenever $k_1 = -k_2$. Westlake (1976) argued that the classic CI needs to be modified so that it is symmetric around 0 for the mean difference since confidence limits are mostly stated in a symmetric form and also to make it easier for non-statisticians to understand.

The CI for $\mu_T - \mu_R$ is given as:

$$-\Delta < \mu_T - \mu_R < \Delta. \tag{3.35}$$

where Δ represents Westlake's symmetric confidence interval.

It is similar to establishing a CI for the T drug formulation (μ_T) being symmetric about the R mean (μ_R), thus :

$$\mu_R - \Delta < \mu_T < \mu_R + \Delta, \tag{3.36}$$

where $\Delta = -k_1\hat{\sigma}_d\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}} + (\bar{Y}_R - \bar{Y}_T)$.

This suggests that

$$(k_1 + k_2)\hat{\sigma}_d\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 2(\bar{Y}_R - \bar{Y}_T). \tag{3.37}$$

The T drug can be considered to be average bioequivalent to the R drug according to the ± 20 rule if $|\Delta| < 0.2\mu_R$. The values of k_1 and k_2 could be obtained by solving Equation 3.38,

$$\int_{k_1}^{k_2} T dt = 1 - 2\alpha, \tag{3.38}$$

for k_1 and k_2 under the constraint conditions in Equation 3.37.

3.7.5 Simulated Annealing Algorithm

Estimation of the AUC after the administration of a drug is one of the challenges encountered in PK studies. During the design phase of bioequivalence studies, it is important to arrange the times of observations in order to estimate accurately the AUC (Westlake, 1979). The choice of the study design affects the efficiency of the PK studies as well as the precision with estimating the parameters. According to Katz and D'Argenio (1983), the observation times should be selected in order to get a minimal expected value of the squared difference between the quadrature approximation and the exact value of the integral. The process of finding the best suitable way of arranging the observation times is a combinatorial optimisation problem which can be achieved

91

using several optimisation algorithms like simulated annealing, tabu search, genetic algorithm, and others (Choi *et al.,* 2007).

A MATLAB program was written to estimate the times, $t_1, t_2, ..., t_m$, that will minimise the objective function using the Simulated Annealing algorithm. The first and last sampling times were fixed and only the interior times were optimised. The process starts by fixing the first and the last sampling points with an iteration process exploring the neighbourhood for the best solution. The best solution with the least objective function value gives the best optimised estimate value of AUC. During the process other parameters of the drug, such as the volume of distribution, fraction of the absorbed dose which is administered, the absorption rate, the elimination rate and the dose that is administered, are also considered. The algorithm stops whenever there are no improved solutions.

The Simulated Annealing recipe:

1. Select the starting value and initial parameter values;

2. Randomly select a new point in the neighbourhood of the original starting value;

3. Compare the two points using the metropolis criterion;

4. Repeat steps 2 and 3 until the system reaches the equilibrium state; and

5. Stop when system reaches optimal solution.

3.8 Population Pharmacokinetics(PopPK) Models

The PopPK model, according to Sheiner *et al.* (1972, 1977), is an addition to the PK model but this time integrates the individual variations. This model is able to quantify the dose response connection using few concentration measurements from each subject in the research as well as other data including weight, age and sex. Bauer *et al.* (2007) stated that new algorithms including NLMEM, the STS and Bayesian hierarchical models have been developed and used for PopPK modelling. The NLMEM is used in this dissertation.

3.9 The Pharmacokinetic Model

The general pharmacokinetic model is:

$$y_{ij} = f(t_{ij}, \phi_i) + g(t_{ij}, \phi_i, \xi_i)\varepsilon_{ij} \quad 1 \le i \le N, \ 1 \le j \le n_i, \qquad (3.39)$$

where $\boldsymbol{\phi_i}$ represents the parameter vector of the structural model, *f*, for each individual i, N is the number of subjects, and n_i is the observation for individual i. The residual error model is given by g which could be dependent on the additional parameter vector $\boldsymbol{\xi_i}$. For simplicity in the notation, given that $\psi_i = (\phi_i, \xi_i)$, where ψ_i are the individual parameters of subject i grouped together, Equation 3.39 can be written as:

$$y_{ij} = f(t_{ij}, \psi_i) + g(t_{ij}, \psi_i)\varepsilon_{ij} \quad 1 \le i \le N, \ 1 \le j \le n_i. \qquad (3.40)$$

The residual errors (ε_{ij}) have a mean of zero and standard deviation of one. From Equation 3.39, $f(t_{ij}, \phi_i)$ and $g(t_{ij}, \phi_i, \xi_i)$ are the conditional mean and standard deviation of y_{ij}.

That is:

$$E(y_{ij}|\psi_i) = f(t_{ij}, \phi_i),$$

$$sd(y_{ij}|\psi_i) = g(t_{ij}, \phi_i, \xi_i).$$

3.9.1 The Structural Model

The first step in a PopPK model development is to discover the structural model that normally describes the data obtained. The structural model is denoted by f in Equation 3.39. For any subject i and the individual vector parameters $\boldsymbol{\psi} = (\phi_i, \xi_i)$, $f(t_{ij}, \phi_i)$ gives the observed variable prediction at time t_{ij}. Assuming that there are no errors ($\varepsilon_{ij}=0$), it is the value that is measured at time t_{ij}. This model predicts the concentration of the drug as a function of time. The most commonly used structural models include 1-, 2- or 3-compartment models with different absorption models (Sun, 2010). This model is usually diagnosed by plotting the predicted values versus the concentration conditioned on time, predicted values for random effects versus concentration conditioned on time, weighted residuals versus time, weighted residuals versus predicted value, predicted value versus concentration conditioned on covariates, and predicted values for random effects versus concentration conditioned on covariates (Sun, 2010). The model could be chosen from literature as a starting point for the modelling process. However, if there is no known structural model to start with, one could start with 1-, 2- and 3-compartment models with different elimination and absorption models and choose the best model using different selection criteria including the Akaike information criterion (AIC), the likelihood ratio test (LRT), the Bayesian information criterion (BIC) and diagnostic plots. A basic structural model is usually one without covariates. The structural model explains the underlying patterns in the data set. In this dissertation, a one-compartment PK model with a first order absorption rate and first order elimination rate is provided in Equation 3.41:

$$C(t,\phi) = \frac{DK_a}{VK_a - Cl}(e^{-(Cl/V)t} - e^{-K_a t}), \qquad (3.41)$$

with the PK parameters $\phi=(K_a, V, Cl)$. The observed concentration is assumed to be normally distributed with a constant error model given as:

$$y_{ij}|\psi_i \sim N(c(t_{ij}, \phi_i), a^2),$$

where $\boldsymbol{\psi_i}=(\phi_i, a)$. The PK parameters (K_{ai}, V_i, Cl_i) are log-normally distributed. This model will be used to demonstrate the estimation methods discussed

in Section 3.12 and the model evaluation methods discussed in Section 3.16.

3.9.2 The Residual Error Model

Although the main aim is to estimate the mean of the population parameters, the variability of the parameters among the different subjects in the population should also be observed (Bonate and Steimer, 2006). That is, the study is interested in how much the value of the obtained parameters vary from one subject to another in the population. According to Bonate and Steimer (2006), random effects are necessary in a NLME model not just because they quantify the population variation but also determine a subject's empirical Bayes estimate (EBE) for parameters. The nature of the data often determines the type of error model which is appropriate for the data. For any stated structural model f, the conditional probability distribution of the observation (y_{ij}) is given by the residual error model. In other words, it is given by the probability distribution of the residual errors (ε_{ij}) and the standard deviation $g((t_{ij}, \phi_i, \xi_i))$. There are different types of residual error models. These include:

- Constant error model: This model assumes that $g(t_{ij}, \phi_i, \xi_i) = a_i$. Then, $\xi_i = a_i$ and the model in Equation 3.40 becomes

$$y_{ij} = f(t_{ij}, \phi_i) + a_i \varepsilon_{ij}. \tag{3.42}$$

- Proportional error model: This model assumes that the standard error of the residual error is proportional to the prediction, i.e., $g(t_{ij}, \phi_i, \xi_i) = b_i f(t_{ij}, \phi_i)$. Then $\xi_i = b_i$ and the model in Equation 3.40 becomes

$$y_{ij} = f(t_{ij}, \phi_i)(1 + b_i \varepsilon_{ij}). \tag{3.43}$$

- Combined error model: This model combines a constant and a proportional model with the assumption that $g(t_{ij}, \phi_i, \xi_i) = a_i + b_i f(t_{ij}, \phi_i)$, $\xi_i = (a_i, b_i)$ and

$$y_{ij} = f(t_{ij}, \phi_i) + (a_i + b_i f(t_{ij}, \phi_i))\varepsilon_{ij}. \tag{3.44}$$

3.9.3 The Covariate Model

A covariate can be described as any variable that is distinct to a subject and could impact how the drug affects the body (Bonate and Steimer, 2006). The covariates of a subject could be categorised as intrinsic factors (genetically determined) such as weight, age, race or extrinsic factors such as dose or smoking status (Bonate and Steimer, 2006). These covariates could be continuous or categorical.

3.10 Parameter Estimation

3.10.1 The Maximum Likelihood Estimation of Population Parameters

This is an approach which is used for estimating the population parameters based on a probability distribution such as the maximum likelihood (ML) estimation (Lavielle, 2014). The ML estimation of the population parametric vector, $\boldsymbol{\theta}$ is made up of maximising, with respect to θ, the observed likelihood function which is defined by:

$$\mathcal{L}_y(\boldsymbol{\theta}) = p(\boldsymbol{y}; \theta) = \int p(\boldsymbol{y}, \boldsymbol{\psi}; \theta) d\boldsymbol{\psi}. \tag{3.45}$$

Given a parameter vector $\boldsymbol{\theta^*}$ where observations were generated by the model parametrised by θ^*. According to Lehmann and Casella (2006), the ML estimation under normal conditions has several attractive properties as the number of individuals in the study increases. These include:

- consistency: the ML estimates converge to the true parameter values;

- asymptotic normality: as the number of subjects in the study increases, the ML estimate distribution approaches the normal distribution with mean θ^* and covariance matrix equal to the inverse of the Fisher information matrix; and

- efficiency: when the sample size approaches infinity, it achieves the Cramer-Rao lower bound implying that no estimator has a lower asymptotic mean squared error than the ML estimator.

One objective in a population modelling method is the estimation of the distribution of the individual parameters $p(\psi_i, \theta)$. For instance, if the vector $\boldsymbol{\psi_i}$ is normally distributed, $\boldsymbol{\psi_i} \sim N(\boldsymbol{\psi_{pop}}, \boldsymbol{\Omega})$, the population parameter vector $\boldsymbol{\psi_{pop}}$ as well as the variance-covariance matrix, $\boldsymbol{\Omega}$ need to be estimated.

In order to calculate the ML estimate, it requires an algorithm to maximise $\int p(\boldsymbol{y}, \boldsymbol{\psi}; \theta) d\boldsymbol{\psi}$ with respect to θ. One algorithm, the stochastic approximation expectation maximisation (SAEM), is used for the estimation. This algorithm has been proven to be efficient and converges well (Kuhn and Lavielle, 2005; Allassonniere et $al.$, 2010).

3.10.2 Expectation Maximisation (EM) Algorithm

An estimator could be determined by maximising the joint distribution $p(y, \psi, \theta)$ if the individual parameters were obtained. According to Dempster et $al.$ (1977), since the individual parameters are not observed, the EM algorithm takes the place of ψ by its conditional expectation. An iteration k updates θ_{k-1}^{EM} to θ_k^{EM} when initial conditions θ_0 are provided in the following two steps:

1. E step: $Q_k^{EM}(\theta) = E(log p(\boldsymbol{y}, \boldsymbol{\psi}; \theta) | \boldsymbol{y}; \theta_{k-1}^{EM}); and$

2. M step: $\theta_k^{EM} = arg\ max\ \ Q_k^{EM}(\theta).$

The EM sequence has the ability to converge under mild conditions to a point of the observed likelihood (Wu, 1983). The relationship between the observation y and the parameter ψ is usually nonlinear which makes the first step not explicit for a nonlinear mixed effect model (Lavielle, 2014). This can be addressed by using the SAEM algorithm as a stochastic approximation which depends on the simulation of ψ (Delyon et $al.$, 1999).

3.10.3 Stochastic Approximation of EM (SAEM) Algorithm

The SAEM algorithm is a stochastic algorithm which is used to estimate the ML estimates (Lavielle, 2014). This algorithm is implemented in a number of softwares including Monolix, NONMEM, MATLAB and R. This is an algorithm based on iterations and requires an initial guess of the PK parameters. The iteration k of SAEM comprises of three stages:

1. the simulation stage:

   for i=1,2,...,N, draw ψ_i^k from the conditional distribution $P(\psi_i|y_i; \theta_{k-1})$;.

2. stochastic approximation:

   Update $Q_{k-1}(\theta)$ according to $Q_k(\theta) = Q_{k-1}(\theta) + \gamma_k(logp(\boldsymbol{y}, \boldsymbol{\psi}^k; \theta) - Q_{k-1}(\theta))$, where γ_k represents a sequence of decreasing positive numbers with $\gamma_1 = 1$; and

3. maximisation step:

   Update θ_{k-1} as stated by $\theta_k = argmax\ Q_k(\theta)$.

For the SAEM algorithm to converge, the requirement is that $\sum_{k=1}^{\infty} \gamma_k = \infty$ and $\sum_{k=1}^{\infty} \gamma_k^2 < \infty$ (Delyon $et\ al.$, 1999). This condition is satisfied if, for instance, γ_k decreases as $1/k$. In order for SAEM to converge, the choice of step-size (γ_k) is very important. Choosing smaller step-sizes ensure there is almost a guarantee of convergence of the algorithm to the ML estimates.

3.11 Model Evaluation

It is very important to evaluate the performance of any model that is developed. In this, the model should be able to explain a phenomenon and the data used in the modelling process. Model evaluation, therefore, is concerned with whether the model best describes the observed data satisfactorily, whether the model is simple enough for extrapolation and whether the model could be used for

the reason it has been developed (Comets *et al.*, 2010). It is essential to check whether the data are in agreement with the model and also if the model explains the data well (Lavielle, 2014). To be able to do this effectively, model diagnostics are used to choose the model that best describes the data and eliminate the models that are not able to reproduce the data (Comets *et al.*, 2010). These diagnostic plots are able to explain whether the model addresses every relevant aspect of the data or if the model needs any further attention. It is also important to be able to conclude if the data best describe a one- or two-compartment model. A model selection process should be developed to select the best model. Whenever several models are valid, it is often desirable to select the model with the simplest assumptions. This selection process is often done using model diagnostics. Another process could be using selection tools to compute a criterion for comparing the models with one another. Some of the criteria which would be used include Akaike information criteria (AIC) and the Bayesian information criteria (BIC).

3.12   Model Diagnostics

The study examines several diagnostic plots and applies them to the data used in this dissertation. These model diagnostics plots would be used to select the best model among the different models which are used. The diagnostic plots which are used in this dissertation are discussed below.

3.12.1   Spaghetti Plot

This is a plot of the concentration time curve for all subjects plotted on the same panel. It is a plot that shows the effects of drugs on subjects after the administration of the drug. This plot can be used to track results of drugs amongst subjects. Spaghetti plot shows variability between individual concentration data at a given time. It gives a better picture of the variability between the individuals in the study. It is easy to select individuals who deviate from the central tendency with respect to half-life, absorption, distribution,

clearance,etc. using this plot.

3.12.2 Individual Fits

For the model defined in Equation 3.40, estimating the individual (ψ_i) and population parameters ψ_{pop} enables the computation for each individual, the predicted profile which is given by the estimated population model ($f(t, \hat{\psi}_{pop})$), and the predicted profile which is given by the estimated individual model, where $\hat{\psi}_i$ is an estimate of ψ_i (Lavielle, 2014).

3.12.3 Observation vs Prediction

The population and individual models enable the calculation of the predictions $f(t_{ij}; \hat{\psi}_{pop})$ for the population and $f(t_{ij}; \hat{\psi})$ for each individual at the observation times t_{ij}.

3.12.4 Residuals

There are a number of residuals but two of them are used in this dissertation. The individual weighted residuals (IWRES) and normalised prediction distribution errors (NPDE).

- IWRES are the estimates of the standardised residual (ε_{ij}) which is based on individual predictions:

$$IWRES_{ij} = \frac{y_{ij} - f(t_{ij}; \hat{\psi})}{g(t_{ij}; \hat{\psi})}.$$

  Whenever the residuals are assumed to be correlated, it can be decorrelated by multiplying each individual vector $\boldsymbol{IWRES_i} = (IWRES_{ij}, \quad 1 \leq j \leq n_i)$ by $\hat{R}_i^{-1/2}$, where $\hat{R}_i$ represents the estimated correlation matrix of the vector of residuals (Lavielle, 2014).

- Population weighted residuals (PWRES) are defined as the normalised difference between observations and their mean. Let $\boldsymbol{y_i} = (y_{ij}, \quad 1 \leq j \leq n_i)$ represent the vector of observations for subject i. Then the

mean of y_i is the vector $\boldsymbol{E(y_i)} = (E(f(t_{ij}; \psi_i)), 1 \leq j \leq n_i)$. Let V_i be the $n_i \times n_i$ variance-covariance matrix of y_i. Then the ith vector of population weighted residuals $\boldsymbol{PWRES_i} = (PWRES_{ij}, \quad 1 \leq j \leq n_i)$ is therefore defined as:

$$PWRES_i = V_i^{-1/2}(y_i - E(y_i)).$$

$E(y_i)$ and V_i are unknown but can be estimated by a Monte Carlo simulation.

- NPDE are a non parametric category of PWRES which depends on rank statistics (Lavielle, 2014). For any (i, j), let $F_{ij} = F_{PWRESij}(PWRES_{ij})$ where $F_{PWRESij}$ is the cumulative distribution function (cdf) of PWRES_{ij}. The NPDE are defined as the empirical estimates of $\Phi^{-1}(F_{ij})$ where the F_{ij} are obtained using Monte Carlo simulation: a large number of replicates $\boldsymbol{y^1, y^2, ..., y^k}$, of the original data $\boldsymbol{y^{obs}}$ are drawn under the model and F_{ij} estimated by:

$$\hat{F}_{ij} = \frac{1}{K} \sum_{k=1}^{K} 1_{y_{ij}^k \leq y_{ij}^{obs}}.$$

The NPDE are then defined as the empirical estimation of $\Phi^{-1}(F_{ij})$, that is, $\mathrm{NPDE}_{ij} = \Phi^{-1}(\hat{F}_{ij})$.

3.12.5 Visual Predictive Checks

This is a diagnostic tool which is used for continuous data which summarises the structural models in the same plot by calculating several quantiles of the empirical distribution of the data used (Lavielle, 2014). Ideally, the visual predictive check (VPC) diagnoses both the fixed and random effects in a mixed-effects model. This is done by comparing different percentiles of the observed data to percentiles of simulated data (Bergstrand *et al.,* 2011). This is often achieved by grouping the data into bins over a range of intervals. The VPC is used to compare different models, suggests if the model needs an improvement and can also be used to support the appropriateness of a model

(Karlsson and Holford, 2008). The prediction intervals for the quantiles are then estimated using the Monte Carlo approach. The model is used to make repeated simulations of the observations using the original design of the data. The percentiles of the simulated data are then plotted versus time from when the treatment started. The same percentiles are then plotted for the observed data to assist in comparing predictions with the observations. The percentiles which are often chosen are the 10th, 50th and 90th percentiles. The percentages of outliers which are outside the prediction interval are estimated to show that the trends in the prediction interval could be used to identify model misspecification (Wilkins *et al.*, 2006).

3.12.6 Model Selection

The statistical tools which are often used to select the best model include the information criteria such as the AIC and the BIC. The AIC and BIC, which are used together with other diagnostic plots were implemented in MONOLIX 4.4.0 and are represented by

$$AIC = -2\mathcal{LL}_y(\theta) + 2P, \tag{3.46}$$

and

$$BIC = -2\mathcal{LL}_y(\theta) + log(N)P, \tag{3.47}$$

where P represents the total number of parameters to be estimated while N is the number of subjects. In selecting the best model between two different models, the model which has the smallest AIC or BIC is chosen as the most appropriate one.

According to Lavielle (2014), when comparing two models, the LRT uses the test statistic

$$LRT = 2(\mathcal{LL}_y(\hat{\theta}_1) - \mathcal{LL}_y(\hat{\theta}_0)),$$

where $\hat{\theta}_0$ and $\hat{\theta}_1$ represent the ML estimates of two models. The distribution of the LRT could either be a χ^2 distribution or a mixture of χ^2 and Dirac delta

(δ) distributions (Lavielle, 2014).

In order to perform the LRTs and compute the information criteria for the models used, it requires the computation of the log-likelihood

$$\mathcal{LL}_y(\hat{\theta}_1) = log(\mathcal{L}_y(\hat{\theta}_1)) = log(p(y; \hat{\theta})),$$

where $\hat{\boldsymbol{\theta}}$ represents the vector of population parameter estimates for the model. The log-likelihood can be estimated for the data using the Monte Carlo approach based on importance sampling. This method provides an unbiased estimate of the log-likelihood whose variance is controlled by the Monte Carlo size. Importance sampling is a sampling tool used for Monte Carlo computing. It refers to a collection of Monte Carlo methods where a mathematical expectation with respect to a target distribution is approximated by a weighted average of random draws from another distribution (Tokdar and Kass, 2010).

3.12.7 Model Validation

Model validation is done to examine whether the final model is a good description of the validation data in terms of the proposed application (FDA, 1999). This can be interpreted as the assessment of the predictability of the developed model to a validation data set. The data which are not used for building the model and parameter estimation is referred to as the validation data. Two model validation types which are often used include external validation and internal validation.

3.12.7.1 External Validation

This is the application of the final developed model on a different dataset which is usually from another study.

3.12.7.2 Internal Validation

This technique involves resampling methods (cross-validation and bootstrapping) and data-splitting (FDA,1999). Data splitting is a convenient and use-

ful internal validation method used to check the predictive performance of a developed model when it is difficult to collect new data to be used for the validation process. Another internal validation method is re-sampling. This process involves cross-validation and bootstrapping. Cross-validation is the use of repeated data-splitting and is very useful because the size of the model development data could be larger than in other validation methods and variability is reduced.

Bootstrapping is another method that is used whenever the sample size is small. It is used for evaluating the performance of a model when there is no validation data set.

3.13 Summary

The Chapter provided a detailed description of the data and the variables that were used in the PopPK model development as identified from previous research. The steps that were taken in building the PopPK model were discussed including the data preparation, discussion of the structural and residual error model as well as the parameter estimation methods. Different model evaluation methods were discussed including the AIC, BIC, individual plots, observation versus prediction plots, plots of residuals, visual predictive checks plots,etc.

Chapter 4

Analysis and Results

4.1 Introduction

As discussed in earlier chapters, bioequivalence studies play an important role in drug development and are conducted primarily by pharmaceutical companies who want market access for their drugs. This chapter discusses the results obtained from the various methods recommended by the FDA and other bodies for bioequivalence studies. The techniques that have been used include ANOVA, CI, TOST, Wilcoxon-Mann-Whitney test and the results from these methods are presented and discussed in this chapter. Results from ABE and PBE are presented and discussed comprehensively.

In addition, this chapter discusses the results of the PopPK model and reviews the model diagnostics generated from the best fitting model. The primary objective is to get the best model that describes the data adequately. The techniques that have been used to assess the different models developed include the AIC, BIC and other diagnostic plots which are illustrated in Figures 4.3 to 4.10. The model validation results together with statistical and diagnostic plots are also provided.

The chapter also presents the results of the SA algorithm which are used to obtain the optimal sampling times for the blood samples to be drawn from the

subjects in the study. The optimised sampling times are then used to generate new concentrations for each subject and bioequivalence testing methods applied to the optimised data.

4.2   ANOVA Results

In a randomised, single-dose, two-period, two sequence COD, it is assumed that the within-subject variances are independent of the formulations resulting in the ANOVA table. The problem of the 2×2 COD concerns unbiased estimation of the formulation differences in the presence of different carryover effects. A failure to detect such an effect may lead to a biased estimate of the formulation difference. Chow and Liu (2008) suggested performing a preliminary test for the presence of different carryover effects before comparing the formulations. Carryover effects are confounded with sequence effects and formulation-by-period interaction. Different sequence effects may not bias the analysis but a difference in carryover effects and a formulation-by-period effect may lead to a serious problem. In bioequivalence studies, carryover effects seldom occur if there is an adequate washout period between the periods. Also, since healthy volunteers are recruited, their physical condition is unlikely to change from one period to another. The pharmacokinetic parameters are subjected to ANOVA in which the variance is partitioned into components due to subjects, periods and formulations. The summary of the results from the log-transformed parameters using ANOVA are presented in Table 4.1:

106

Table 4.1: ANOVA table for log-transformed AUC_{0-t}.

Source of Variation	DF	SS	MS	F statistics	P-value
Between:Subject					
Sequence	1	0.1053	0.1053	2.86	0.1129
Subject(Sequence)	14	0.5156	0.0368	3.25	0.0176
Within:Within					
Formulation	1	0.0028	0.0028	0.25	0.6260
Period	1	0.0572	0.0572	5.04	0.0415
Residuals	14	0.1589	0.01135		
Total	31	0.8398			

From Table 4.1, the results indicate no significant effect in the sequence (p-value=0.1129) but has significant subject effect (p-value=0.0176) at the 5% level of significance. The significant subject effect therefore suggests that there are inter-subject dissimilarity in clearance and also in the AUC_{0-t} as well. In addition, the period effect is also significant (p-value=0.0415) and the formulation effect (p-value=0.6260) is not significant at the 5% level of significance. This suggests that there are no statistically significant carryover effect or formulation effects in this bioequivalence study. A significant period effect does not suggest bioinequivalence. The formulation effect test is a secondary test for the equality of the predicted means against the expected mean difference. That is:

$$H_0 : exp(\mu_T) = exp(\mu_R) \quad vs. \quad H_1 : exp(\mu_T) \neq exp(\mu_R).$$

It is therefore not appropriate to use the formulation effect test for bioequivalence assessment even if it is significant.

The residual variance estimate $\hat{\sigma}_W^2 = MS_{within} = 0.011348$. This variability is better expressed as a coefficient of variation (CV) for easy interpretation. The residual variance estimate for the AUC_{0-t} is 0.011348 ($\hat{\sigma}_W^2 = 0.011348$)

and is proportionate to the within-subject CV of 10.7% assuming that the formulation variances for the within-subject are equal. Similarly, the between-subject CV is 11.3%.

Table 4.2: ANOVA table for log-transformed $AUC_{0-\infty}$.

Source of Variation	DF	SS	MS	F statistics	P-value
Between:Subject					
Sequence	1	0.0615	0.0615	2.12	0.1673
Subject(Sequence)	14	0.4060	0.0290	2.06	0.0938
Within:Within					
Formulation	1	0.0149	0.0149	1.06	0.3199
Period	1	0.0407	0.0407	2.90	0.1108
Residuals	14	0.1967	0.0140		
Total	31	0.7198			

Table 4.2 provides the results for the ANOVA of the log-transformed extrapolated AUC ($AUC_{0-\infty}$). The result indicates no significant effect in the sequence (p-value=0.1673) and also no significant subject effect (p-value=0.0938) at the 5% level of significance. The non-significant subject effect therefore suggests that there are no inter-subject differences in the $AUC_{0-\infty}$ and the clearance. In addition, the period effect is not significant (p-value=0.1108) and the formulation effect (p-value=0.3199) as well is not significant at the 5% level of significance. This suggests that there are no statistically significant carryover effect, period effect or formulation effects in this bioequivalence study.

The residual variance estimate $\hat{\sigma}_W^2 = MS_{within} = 0.014049$. This variability is better expressed as a CV for easy interpretation. The residual variance approximation for $AUC_{0-\infty}$ is 0.014049 ($\hat{\sigma}_W^2 = 0.014049$) and this is equivalent to the within-subject CV of 11.9% assuming that the formulation variances for the within-subject are equal. Similarly, the between-subject CV is 8.7%. Table 4.3 provides the results for the ANOVA of the log-transformed C_{max}.

Table 4.3: ANOVA table for log-transformed C_{max}.

Source of Variation	DF	SS	MS	F statistics	P-value
Between:Subject					
Sequence	1	0.6110	0.6110	2.84	0.1140
Subject(Sequence)	14	3.0108	0.2151	1.04	0.4706
Within:Within					
Formulation	1	0.0760	0.0760	0.37	0.5538
Period	1	0.0029	0.0029	0.01	0.9072
Residuals	14	2.8924	0.2066		
Total	31	6.5932			

The results indicate no significant effect in the sequence (p-value= 0.1140) and the subject effect is also not significant (p-value= 0.4706) at a 5% level of significance. The non-significant subject effect therefore suggests that there are no inter-subject contrasts in clearance and in the C_{max} as well. To add, there is a non-significant period effect (p-value= 0.9072) while the formulation effect is also not significant (p-value=0.5538). This suggests that there are no statistically significant carryover effect, period effect or formulation effects in this bioequivalence study.

The residual variance estimate $\hat{\sigma}_W^2 = MS_{within} = 0.2066$. This variability is better expressed as a CV for easy interpretation. The residual variance evaluation for C_{max} is 0.014049 ($\hat{\sigma}_W^2 = 0.014049$) and this is comparable to the within-subject CV of 47.9% assuming that the formulation variances for the within-subject are equal. Similarly, the between-subject CV is 6.5%.

4.3 Statistical Methods for Average Bioequivalence

4.3.1 Confidence Interval Method

The common approach for bioequivalence testing as stated by the FDA is to construct a 90% CI for the mean difference of the log-transformed responses of the PK parameters; AUC_{0-t}, $AUC_{0-\infty}$ and C_{max}. In order to conclude that the R and T drugs are bioequivalent, the requirement states that the confidence interval must be within the range [0.8-1.25].

4.3.1.1 The Classic Confidence Interval

Table 4.4: The mean of the different log-transformed pharmacokinetic parameters.

	Mean	
Parameter	Test	Reference
AUC_{0-t}	4.115988	4.134758
$AUC_{0-\infty}$	4.197142	4.240359
C_{max}	1.237731	1.140246

Table 4.5: Confidence Interval for the different pharmacokinetic parameters.

	AUC_{0-t}		$AUC_{0-\infty}$		C_{max}	
	Interval Limits					
	Lower	Upper	Lower	Upper	Lower	Upper
Shortest	98.32	100.77	97.65	100.31	89.59	127.51

Table 4.4 presents the mean of the different log-transformed R and T treatment for each of the pharmacokinetic parameters. Table 4.5 provides the confidence intervals for the different log-transformed pharmacokinetic parameters.

The mean AUC_{0-t} values were 4.115988 and 4.134758 for the T and R drugs, respectively. The mean $AUC_{0-\infty}$ for the T and R drugs were 4.197142 and 4.240359, respectively. The mean C_{max} values were 1.237731 and 1.140246 for the T and R drugs, respectively. The mean AUC_{0-t} and $AUC_{0-\infty}$ for the R drug is higher than the T drug mean for the same parameters. However, the mean C_{max} for the T drug is higher than that of the R drug. The 90% confidence interval for AUC_{0-t} and $AUC_{0-\infty}$ were within the 80-125% limit set by FDA and other regulatory bodies. However, the 90% confidence interval for the C_{max} was not within the 80-125% limit.

Table 4.6 presents the 90% confidence limits and the decisions reached about bioequivalence between the R and T drugs.

Table 4.6: Confidence Interval and the decision about bioequivalence.

	Interval Limits		
Parameter	Lower	Upper	Decision
AUC_{0-t}	98.32	100.77	ABE claimed
$AUC_{0-\infty}$	97.65	100.31	ABE claimed
C_{max}	89.59	127.51	ABE not claimed

The 90% confidence interval for the log-transformed AUC_{0-t} is presented as (98.321,100.771). This confidence interval is within the (80-125) range according to the FDA guidelines. Therefore the R drug and the T drug are declared as average bioequivalent with regards to the AUC_{0-t}.

The associated 90% confidence interval for the $AUC_{0-\infty}$ as presented in Table 4.6 is (97.65,100.31) and is well within the interval (80-125) range. The T and R formulations are therefore average bioequivalent with regards to the $AUC_{0-\infty}$.

The associated 90% confidence interval for C_{max} is (89.59,127.51) which is not contained within the (80-125) range as per the FDA guidelines. Average

bioequivalence cannot be concluded as required by the FDA guideliness as the upper confidence limit exceeds the required upper limit as per FDA guidelines. However, for C_{max}, the WHO, the SAMCC and the EMEA adopt a more relaxed equivalence interval of (70-133) for C_{max}. By virtue of this relaxed margin used by these organisations, bioequivalence can be claimed.

4.3.1.2 Westlake's Symmetric Confidence Interval

For any two formulations to be declared as average bioequivalent using the Westlake symmetric confidence interval, Δ should be less than $0.2\mu_R$, where Δ is defined in Equation 3.36. From Table 4.7, the Δ values for AUC_{0-t} and $AUC_{0-\infty}$ are less than $0.2\mu_R$ while the Δ value for C_{max} is greater than $0.2\mu_R$. Therefore, average bioequivalence is concluded in terms of AUC_{0-t} and $AUC_{0-\infty}$ but not C_{max}.

Table 4.7: Westlake's CI results for log-transformed AUC_{0-t}, $AUC_{0-\infty}$ and C_{max}.

Parameter	Westlake's Δ value	$0.2\mu_R$
AUC_{0-t}	0.07307435	0.8269516
$AUC_{0-\infty}$	0.1001362	0.8480718
C_{max}	0.3242763	0.228048

4.3.1.3 Schuirmann's TOST Approach

The TOST procedure was introduced by Schuirmann (1987) and is based on Equation 3.13. Two formulations are considered as bioequivalent if H_{01} and H_{02} from Equation 3.13 are both rejected at a significance level which is pre-determined.

The TOST results for the following log-transformed PK parameters AUC_{0-t}, $AUC_{0-\infty}$ and C_{max} are presented in Table 4.8.

Table 4.8: The Schuirmann's TOST results.

	T_L	T_U	$t(\alpha, n_1 + n_2 - 2)$
AUC_{0-t}	21.458	-22.455	1.761
$AUC_{0-\infty}$	19.206	-21.269	1.761
C_{max}	2.026	-0.812	1.761

From the results, $|T_L|$ and $|T_U|$ for AUC_{0-t} and $AUC_{0-\infty}$ are both greater than $t(0.05, 14) = 1.761$. Since these values are greater than 1.761, the null hypotheses (H_{01} and H_{02}) in Equation 3.13 are both rejected at 5% level of significance. Bioequivalence can be concluded according to the ± 20 rule for both AUC_{0-t} and $AUC_{0-\infty}$. However, for C_{max}, $|T_L|$ is greater than 1.761 but $|T_U|$ is less than 1.761. Bioequivalence cannot be claimed using the C_{max}.

4.3.2 Non-Parametric TOST Method

4.3.2.1 Wilcoxon-Mann-Whitney Two One-Sided Test for T_{max}

Table 4.9 lists the ranks of b_{hik}.

Table 4.9: Ranks of b_{hik} for data.

Sequence	Formulation	Subject	Period1	Period2	Subject Total	PD	$2 \times b_{Lik}$	$R(b_{Lik})$	$2 \times b_{Uik}$	$R(b_{Uik})$
1	RT	842	62.4935	53.378	115.8715	-9.1155	16.18328	9	-34.4143	2
1	RT	858	60.244	63.4214	123.6654	3.1774	28.47618	11	-22.1214	3
1	RT	860	65.0565	78.4096	143.4661	13.3531	38.65188	14	-11.9457	6
1	RT	861	50.333	56.0545	106.3875	5.7215	31.02028	13	-19.5773	5
1	RT	862	49.2165	38.5985	87.815	-10.618	14.68078	8	-35.9168	1
1	RT	864	60.549	78.99	139.539	18.441	43.73978	16	-6.85777	9
1	RT	869	57.7215	71.838	129.5595	14.1165	39.41528	15	-11.1823	7
1	RT	870	49.312	54.3085	103.6205	4.9965	30.29528	12	-20.3023	4
2	TR	827	56.264	55.291	111.555	-0.973	-0.973	2	-0.973	10
2	TR	843	61.1015	73.1955	134.297	12.094	12.094	6	12.094	14
2	TR	844	59.2165	72.3501	131.5666	13.1336	13.1336	7	13.1336	15
2	TR	857	60.798	66.7527	127.5507	5.9547	5.9547	3	5.9547	11
2	TR	859	68.9234	85.6995	154.6229	16.7761	16.7761	10	16.7761	16
2	TR	863	67.7175	77.544	145.2615	9.8265	9.8265	5	9.8265	13
2	TR	866	66.153	59.1085	125.2615	-7.0445	-7.0445	1	-7.0445	8
2	TR	985	58.9715	67.0836	126.0551	8.1121	8.1121	4	8.1121	12

From Table 4.9, the value of R_L and R_U are found to be 98 and 37, respectively.

$$W_L = R_L - \frac{n_1(n_1 + 1)}{2},$$
$$= 98 - \frac{8(8 + 1)}{2},$$
$$= 62,$$

and

$$W_U = R_U - \frac{n_1(n_1 + 1)}{2},$$
$$= 37 - \frac{8(8 + 1)}{2},$$
$$= 1.$$

Therefore $W(0.95) = 8(8) - 16 = 48$ and $W(0.05) = 16$. Since $W_L = 62$ is greater than $W(0.95) = 48$ and $W_U = 1$ is less than $W(0.05) = 16$, both two one-sided null hypotheses in Equation 3.19 are rejected at 5% level of significance. Average bioequivalence can then be concluded for both formulations using the T_{max}.

4.3.3 Power and Sample Size Determination

Bioequivalence studies should be designed properly such that the correct number of subjects are used to address the objective of the study. When designing clinical trials, a power of at least 80% is required. The required number of samples is determined from the design phase for a specific power. The empirical power estimates for specific sample sizes and their associated CVs are presented in Table 4.10. It is evident that, for any sample size, the power of TOST increases as the CV decreases and the power decreases as the CV increases.

Table 4.10: Values for the index of sensitivity.

	Coefficient of variation				
N	10	15	20	25	30
8	0.916	0.600	0.300	0.133	0.059
10	0.969	0.742	0.442	0.216	0.096
12	0.988	0.831	0.565	0.314	0.151
14	0.996	0.887	0.664	0.417	0.217
16	0.999	0.926	0.737	0.506	0.293

Figures 4.1 and 4.2 present a graphical representation of the sample size and power of the TOST used in bioequivalence studies. The graphs indicate that, as the sample size increases, so does the power of the TOST.
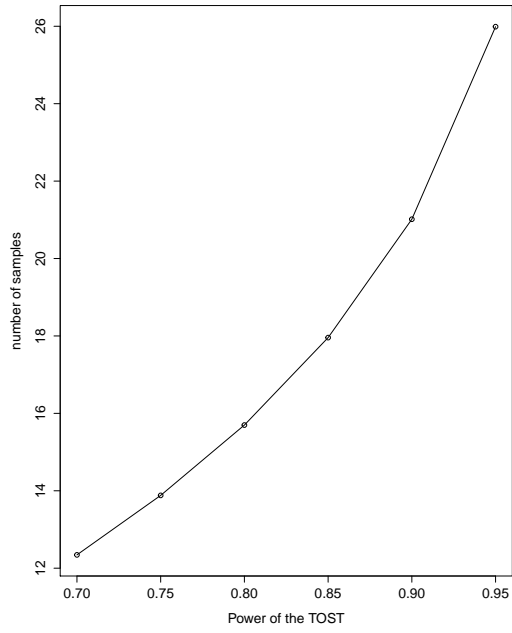
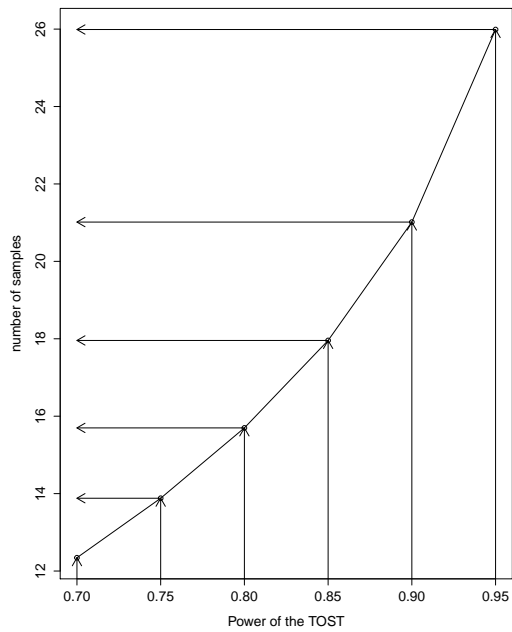Figure 4.1: The power of the TOST with corresponding number of samples.

Figure 4.2: The power of the TOST with corresponding number of samples.

4.4 Population Bioequivalence

For population bioequivalence to be established, Θ_{pop} in Equation 2.7 must be less than the regulatory set value of 1.74483 for the log-transformed AUC_{0-t}, $AUC_{0-\infty}$ and C_{max}. Table 4.11 presents the results of the PBE for three PK parameters.

Table 4.11: PBE results for $logAUC_{0-t}$, $logAUC_{0-\infty}$ and $logC_{max}$.

Parameter	Θ_{pop}	Decision
AUC_{0-t}	1.107772	PBE claimed
$AUC_{0-\infty}$	1.010149	PBE claimed
C_{max}	1.523058	PBE claimed

From Table 4.11, the Θ_{pop} values for all the PK parameters are less than the 1.74 as per the FDA guidelines. This indicates that the two drugs are population bioequivalent, implying that a patient can have the option to be prescribed either the R or the T drug if they have not been using either of the two.

4.5 Individual Bioequivalence

As stated by various researchers including Jones *et al.* (1999), IBE is only conducted for a replicated design so that the within- and between-subject variability is estimated. IBE enables patients to easily switch from one formulation to another formulation without any therapeutic effects. It is therefore impossible to perform IBE in this research as the data are a 2×2 COD.

4.6 Population Pharmacokinetics

A randomised, single dose, two sequences, two periods crossover study was performed on the subjects to compare the PK parameters of the T and the R drug. The treatment periods were well separated and therefore there is no evidence of carryover effect. The data were analysed with a NLMEM using the SAEM algorithm which is implemented in MONOLIX 4.4.0. The NLMEM parameters were estimated using the SAEM algorithm implemented in MONO-LIX. Two different one-compartment models with first order absorption and first order elimination were used to describe the data. Different residual error models were used together with the structural model and the models were compared using BIC, AIC and VPC to determine the best model fit. The ideal statistical model corresponds to the one with the smallest BIC and AIC and also one with a VPC which best represents the data. After fitting the model to the data, different graphs such as individual fits graphs, observation versus predictions graphs, normalised prediction distribution errors (NPDE) graphs, VPC, distribution of the individual parameter graphs, random effects joint distribution graphs and the SAEM convergence graphs were obtained for each of the different models. The best model based, on the criteria mentioned previously, is presented with the model diagnostics and analysed. The VPC graphs were evaluated using the prediction intervals (PI) for each model.

The best statistical model for the data is the combined error model with two parameters a for the additive and b for the proportional part together with the structural model with PK parameters K_a, V and Cl. The model used has no covariates and the period or sequence effects are also not incorporated. Table 4.12 displays parameter estimates with their standard errors and root square errors. Figures 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10 present the spaghetti plot, the individual and population fits, the prediction vs observation plot, the residuals, the VPC plots, the parameter distribution plot, the random effect joint distribution plot and the SAEM convergence plot.

The pharmacokinetic population modelling was done with MONOLIX version 4.4.0. The parameters of the population were estimated using the SAEM algorithm and reported in Table 4.12. These parameter values are considered to be the EBE because they were obtained at the maximum. The relative standard error (r.s.e) percentage for the parameters K_a, V and Cl are 19%, 10% and 5%, respectively. The residual variability was modelled using different error models (constant and combined error model). The best error model was the combined error model for the residual variability using model diagnostics. The AIC, BIC and log-likelihood of the best model are presented in Table 4.13. The interindividual variability (IIV) for the random effects ω_{K_a}, ω_V and ω_{Cl} are 39%, 28% and 13%, respectively. The individual parameters follow a log-normal distribution and are independent as illustrated below:

$$Cl_i = 0.234 \times exp(\eta_{i,Cl}) \qquad with \quad \eta_{i,Cl} \sim N(0, 0.136^2), \qquad (4.1)$$

$$V_i = 5.97 \times exp(\eta_{i,V}) \qquad with \quad \eta_{i,V} \sim N(0, 0.281^2), \qquad (4.2)$$

$$(K_a)_i = 2.76 \times exp(\eta_{i,K_a}) \qquad with \quad \eta_{i,K_a} \sim N(0, 0.391^2). \qquad (4.3)$$

Table 4.12: Parameter estimates using the combined error model.

Parameter	Estimate	s.e (stochastic approximation)	r.s.e (%)
K_a	2.76	0.52	19
V	5.97	0.61	10
Cl	0.234	0.013	5
Omega_K_a	0.391	0.18	46
Omega_V	0.281	0.073	26
Omega_Cl	0.136	0.043	32
a	0.0336	0.0089	27
b	0.164	0.014	9

Table 4.13: Statistical tools for model selection.

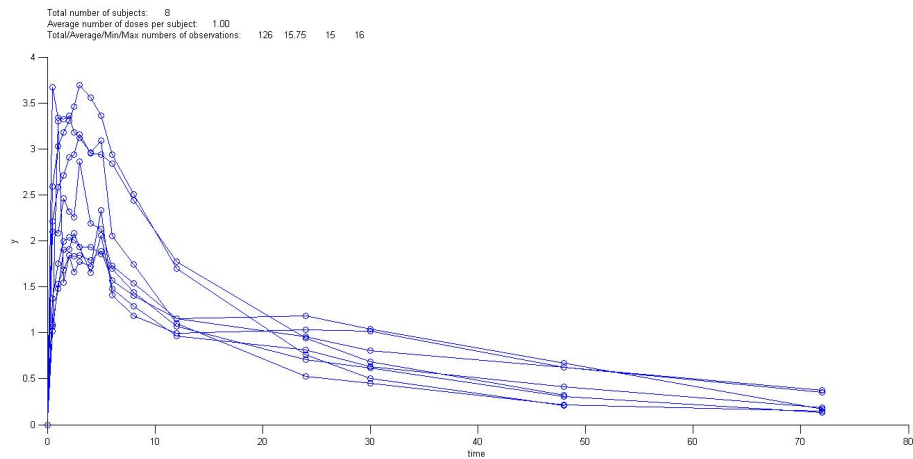| AIC	BIC	-2 × log-likelihood
33.28	33.92	17.28

Figure 4.3: Spaghetti plot with parameters *Ka*, *V* and *Cl*.

Figure 4.3 shows the concentration time graph (spaghetti plot) for the R drug with 8 subjects. It is a plot of the concentration time curve for all subjects plotted on the same panel. This graph represents the distribution of drug absorption, distribution and elimination for the subjects that received the same dose of the drug at the same sampling times. From the graph, it is concluded that almost all the subjects exhibited similar drug kinetics (absorption, distribution, metabolism, elimination). It can also be concluded from the plot that the data do not contain experimental errors and no large inter-individual differences exist with respect to the peak.
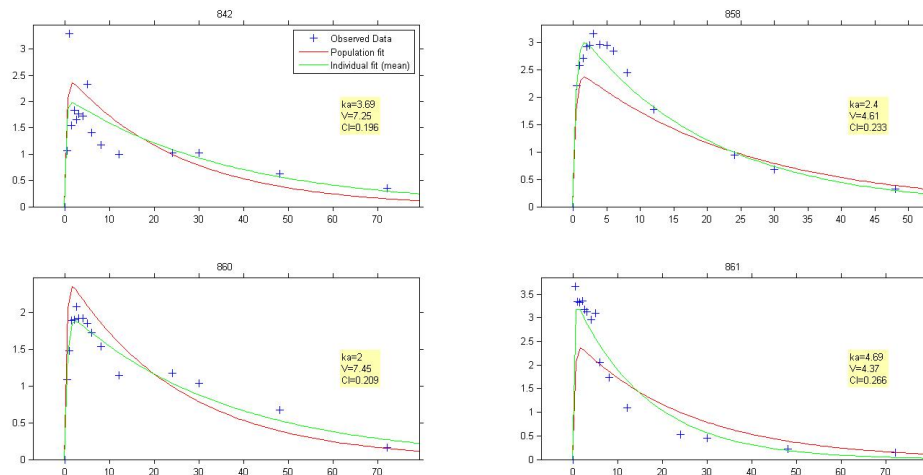
Figure 4.4: The Individual and the population fits for the subjects are plotted simultaneously on each of these four plots.

Figure 4.4 shows the individual fits and the population fits for subjects with their estimated individual parameters (lighter line) and the estimated population parameters (darker line). Also, the estimated individual parameters for each subject are indicated. The data are displayed using the crosses. The population fits is the estimation of the model on each subject based on the population parameters and the individual fits is the estimation on each subject based on the individual parameters. The individual parameters are estimated from the conditional mode or the conditional mean. The individual fits graphs fit the data well and therefore the proposed PK model is accepted since the fits seem acceptable. These fits are acceptable for the one-compartment model which was proposed since the data fit the model.
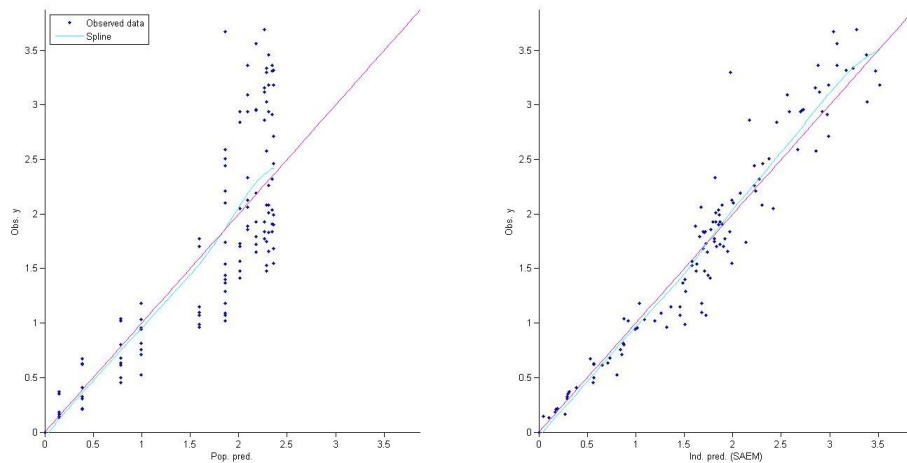
Figure 4.5: Observations versus predictions graphs using the population model(left) and individual model (right).

Figure 4.5 shows the observation versus prediction using the proposed population model (left) with population parameters and the individual model (right) with the individual parameters. The observation versus individual prediction graph indicates that there is no misspecification in the model but the observation versus population prediction graph indicates a slight model misspecification. The data are nicely distributed around the identity line for the individual predictions. There is a slight deviation away from the line for the population prediction as indicated in Figure 4.5. It is also evident from the population prediction graph that there is an inter-individual variability and residual variability since not all the data points lie on the identity line. However, the inter-individual variability is small. There is no nonzero prediction which corresponds to a zero concentration observation which indicates that there is no delay between the time when the drug was administered and the time during which it was absorbed. In other words, there is no lag-time. The amplitude of the residual errors also seem to increase along the predicted concentration for both graphs. This indicates that a constant error model is not the most appropriate model for this type of data.
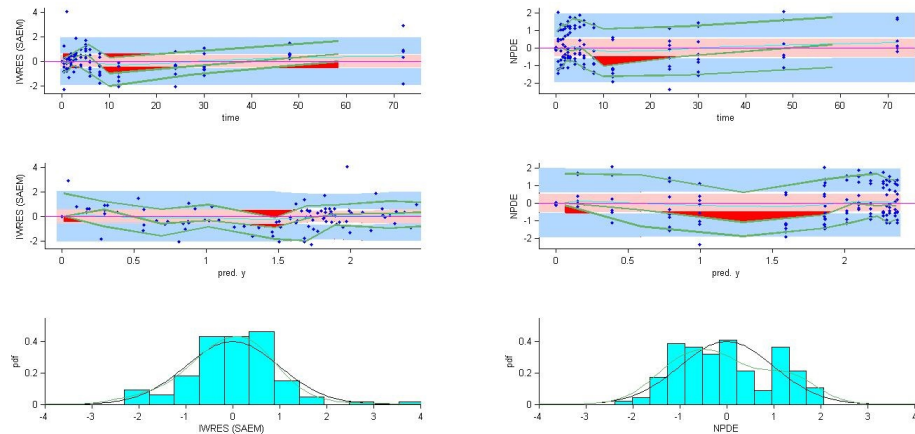
Figure 4.6: The IWRES and the NPDE.

From Figure 4.6, the graphs for the IWRES and the NPDE with respect to time are presented in the top row and the graphs for the IWRES and NPDE with respect to the prediction are in the middle row as well as the graphs for presenting the comparison of the empirical and theoretical probability density function (pdf) of the IWRES and the NPDE are at the bottom. The prediction intervals are also added on each of the graphs. The IWRES and the NPDE are expected to be independent standardised normal random variables. The IWRES and the NPDE histograms also suggest normality for both residuals. From Figure 4.6, both the IWRES and the NPDE suggest that there is a slight model misspecification as indicated by the residuals which are outside the prediction intervals although generally the model is acceptable. The IWRES are centered around the zero mean line indicating that the one compartment structural model is the appropriate model for this data. The variance of the IWRES is constant which also indicates that the error model is appropriate. The IWRES can then be said to be normally distributed with a mean of zero and a constant variance. The amplitude of the residuals are not too large. The $10th$, $50th$ and $90th$ percentiles which are displayed in Figure 4.6 indicate a slight misspecification although generally the model predicts the data quite well. The plot of the NPDE versus predictions tends to disperse the data quite
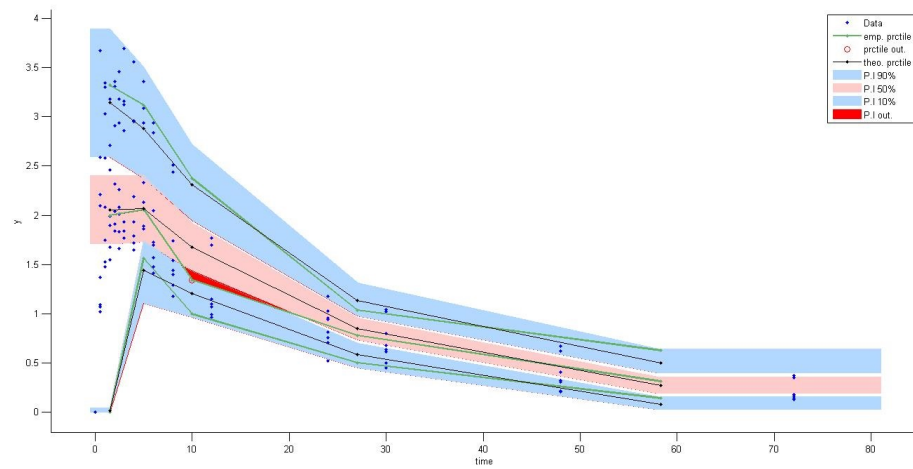
well over the range of predictions.

Figure 4.7: Visual predictive check.

From Figure 4.7, there is an indication that the structural model is good except for between 10 to 19 hours where the model underestimates or under predicts. In the 10% prediction intervals, the $10th$ empirical percentile deviates from the theoretical percentile slightly between 9 to 30 hours although it does not indicate model misspecification. The 90% prediction interval fits the data quite well.
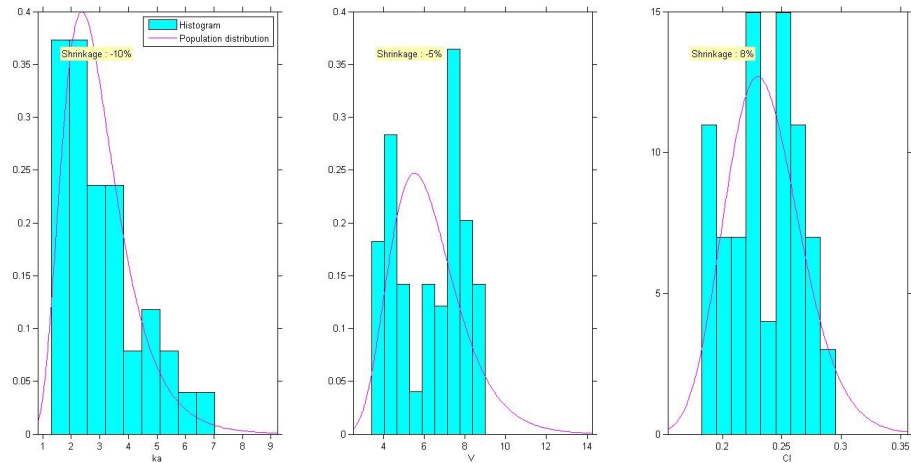
Figure 4.8: Estimated population distributions of the individual parameters (solid line) and the empirical distributions of the individual parameters simulated with their conditional distributions (histogram).

Figure 4.8 indicates the probability density function which was obtained for each individual parameter in the model from the estimated population parameters as well as the empirical distribution which is indicated by the histogram of the individual parameters. Each of the individual parameters was simulated with its own conditional distribution. From Figure 4.8 and from the plots for Ka, V and Cl, it is clear that that the individual parameters are log-normally distributed and the shrinkage which was estimated using the simulated parameters, is a random variable with a mean of zero and variance close to $1/N$. The shrinkage values (-10%, -5% and 8%) as shown in Figure 4.8 are significantly different from zero.
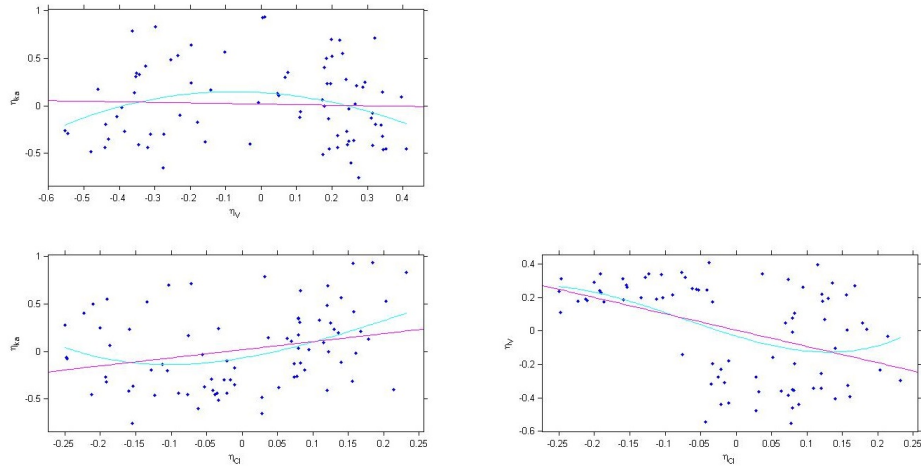
Figure 4.9: Joint distribution of the random effects.

The graphs in Figure 4.9 represent a plot of each pair of simulated random effects against each other. Regression and spline interpolations are also displayed to assist in determining the correlation between the random effects. From the three plots, it is evident that the points are randomly scattered and indicate no trend. This indicates that there is no correlation between the random effects which indicates that there is a diagonal variance-covariance matrix. Therefore K_a, V and Cl are independent of each other.
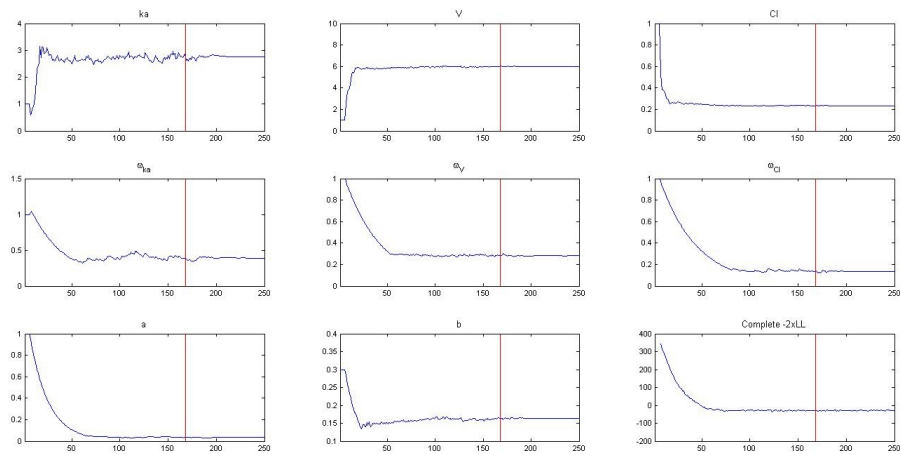
Figure 4.10: SAEM convergence for each of the parameters K_a, V and Cl.

Figure 4.10 shows the parameter estimates which were computed after each iteration of SAEM. It shows the convergence of the estimated parameters. The parameters of the one-compartment model with first-order absorption and linear elimination are estimated. The vertical lines on each graph indicate where the algorithm switches from the first phase to the second phase. The SAEM algorithm converges quite well to the global maximum of the likelihood for each of the parameters.

4.7 Model Validation

The fitting of the PopPK model on a portion of the data as well as the estimated parameters and model diagnostics were discussed in the previous section. Even though the model diagnostics suggested that the one compartment model which was fitted on the data described it well, it does not suggest that the model is highly significant. The developed model has to be validated. The validation of the final PopPK model was based on statistical and graphical methods. The parameter estimates are presented in Table 4.14 and the AIC and BIC values of the validated model are shown in Table 4.15. The model diagnostics plots are also presented in Appendix B.

Table 4.14: Validation model parameter estimates using the combined error model.

Parameter	Estimate	s.e (stochastic approximation)	r.s.e (%)
K_a	1.96	0.24	12
V	3.33	0.8	24
Cl	0.227	0.0058	3
Omega_K_a	0.303	0.098	32
Omega_V	0.675	0.17	25
Omega_Cl	0.0614	0.021	34
a	0.0424	0.01	24
b	0.0859	0.0082	10

Table 4.15: Statistical tools for model selection.

AIC	BIC	-2 × log-likelihood
20.53	21.17	4.53

The AIC and BIC values of the validation model in Table 4.15 suggest that the one compartment model fits the validation data well. The diagnostic plots

in Figures B.2, B.3, B.4, B.5, B.6, B.7 and B.8 which are shown in Appendix B also confirm that the model fits the data.

4.8 Simulated Annealing

In certain cases, the individual PK parameters are either not available or cannot be determined directly. In such situations, the PK parameters are estimated using modelling. The PK parameters which were estimated with modelling are used in the simulated annealing algorithm to generate optimised sampling times. The optimisation approach used in the dissertation uses a one-compartment model with first order absorption as illustrated in Equation 2.19. Mostly, the K_a, K_e and V are random variables that change for every subject but the dose is a fixed and known value. The optimal sampling times which were obtained from the one-compartment model are t=(0,0.5,0.71,1.3,1.44,2.10,5.13,5.16,5.36,7.15,8,19.44,24,33.62,34.53,48,72) and shown in Figure 4.11.
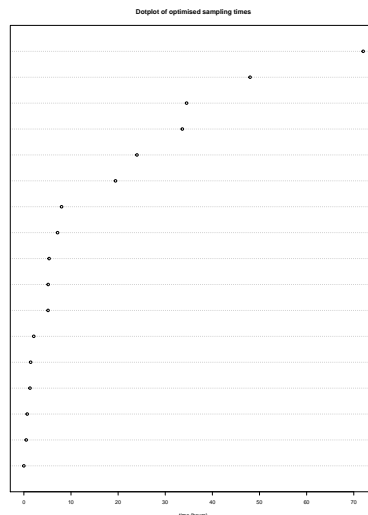

Figure 4.11: Dot plot of optimised sampling times.

These sampling times were obtained from a one-compartment model with population pharmacokinetic parameters $K_a = 0.0393$, $K_e = 2.71$, $D = 15$,

$V = 0.0863$ and $F = 1$. The parameters have been defined in Equation 2.1. The concentrations for each subject, the AUC_{0-t}, the $AUC_{0-\infty}$ were obtained and are presented in Tables 4.16 and 4.17.

Table 4.16: Log-transformed AUC for the T and R formulations after optimisation.

Subject	Sequence	Period	Formulation	AUC_{0-t}	$AUC_{0-\infty}$	C_{max}
842	1	1	R	67.43587	79.30902	2.002654
858	1	1	R	64.62041	66.45337	2.990088
860	1	1	R	63.25278	72.97602	1.908576
861	1	1	R	53.91614	54.42014	3.205287
862	1	1	R	51.74834	56.07939	1.860115
864	1	1	R	62.15424	75.20628	1.732339
869	1	1	R	60.51894	61.10355	3.475313
870	1	1	R	51.95971	53.83509	2.322062
842	1	2	T	65.86338	69.09776	2.66777
858	1	2	T	61.14068	64.37372	2.43851
860	1	2	T	71.41448	73.86414	3.13272
861	1	2	T	58.21259	59.55591	2.813054
862	1	2	T	44.36431	48.88173	1.519091
864	1	2	T	72.333	77.75346	2.631034
869	1	2	T	63.45344	73.11628	1.922489
870	1	2	T	63.17753	67.89929	2.312708
827	2	1	T	58.80359	58.81331	5.714009
843	2	1	T	60.86652	62.05818	3.003115
844	2	1	T	61.41677	69.94059	1.919448
857	2	1	T	65.92797	68.02171	2.953064
859	2	1	T	67.94529	67.94547	8.545359
863	2	1	T	68.45833	68.45838	9.849732
866	2	1	T	66.2853	71.80288	2.366139

Table 4.17: Log-transformed AUC for the T and R formulations after optimisation.

Subject	Sequence	Period	Formulation	AUC_{0-t}	$AUC_{0-\infty}$	C_{max}
985	2	1	T	65.02689	70.92395	2.270761
827	2	2	R	58.21236	71.60807	1.565088
843	2	2	R	72.3221	73.32932	3.794461
844	2	2	R	70.19283	72.04327	3.239377
857	2	2	R	71.18192	71.18194	10.3222
859	2	2	R	88.02288	93.19436	3.379288
863	2	2	R	73.97729	81.16862	3.126044
866	2	2	R	61.17458	75.12056	1.647895
985	2	2	R	66.62649	70.44845	2.580378

The AUC data obtained after the optimisation is used to test for BE studies and the results are presented. The summary of the results from the log-transformed parameters using ANOVA are presented in Table 4.18.

Table 4.18: ANOVA table for the log-transformed AUC_{0-t}.

Source of Variation	DF	SS	MS	F statistics	P-value
Between:Subject					
Sequence	1	0.0812	0.0812	3.98	0.0659
Subject(Sequence)	14	0.2855	0.02039	3.25	0.0174
Within:Within					
Formulation	1	0.00269	0.00269	0.43	0.5230
Period	1	0.0320	0.0320	5.11	0.0402
Residuals	14	0.0878	0.0063		
Total	31	0.4892			

From 4.18, the results indicate no significant effect in the sequence (p-value = 0.0659) but has significant subject effect (p-value=0.0174) at the 5% level of significance. The significant subject effect therefore suggests that there are inter-subject dissimilarity in clearance and also in the AUC_{0-t} as well. In addition, the period effect is also significant (p-value=0.0402) and the formulation effect (p-value=0.5230) is not significant at the 5% level of significance. This suggests that there are no statistically significant carryover effect, period effect or formulation effects in this bioequivalence study. The formulation effect test is a secondary test for the equality of the predicted means against the expected mean difference. That is,

$$H_0 : exp(\mu_T) = exp(\mu_R) \quad vs. \quad H_1 : exp(\mu_T) \neq exp(\mu_R).$$

It is therefore not appropriate to use the formulation effect test for bioequivalence assessment even if it is significant.

The residual variance estimate $\hat{\sigma}_W^2 = MS_{within} = 0.0063$. This variability is better expressed as a CV for easy interpretation. The residual variance estimate for the AUC_{0-t} is 0.0063 ($\hat{\sigma}_W^2 = 0.0063$) and is proportionate to the within-subject CV of 7.92% assuming that the formulation variances for the within-subject are equal. Similarly, the between-subject CV is 6.6%.

Table 4.19: ANOVA table for the log-transformed $AUC_{0-\infty}$.

Source of Variation	DF	SS	MS	F statistics	P-value
Between:Subject					
Sequence	1	0.0630	0.0630	2.83	0.1149
Subject(Sequence)	14	0.312	0.022	2.95	0.0260
Within:Within					
Formulation	1	0.0163	0.0163	2.16	0.1634
Period	1	0.0453	0.0453	5.99	0.0282
Residuals	14	0.1058	0.00756		
Total	31	0.5424			

Table 4.19 gives the ANOVA results of the log-transformed $AUC_{0-\infty}$. The results indicate no significant effect in the sequence (p-value=0.1149) and also a significant subject effect (p-value=0.0260) at the 5% level of significance. The significant subject effect therefore suggests that there are inter-subject dissimilarity in clearance and also in the $AUC_{0-\infty}$ as well. In addition, the period effect is significant (p-value=0.0282) and the formulation effect (p-value=0.1634) is not significant at the 5% level of significance. This suggests that there are no statistically significant carryover effect, period effect or formulation effects in this bioequivalence study.

The residual variance estimate $\hat{\sigma}_W^2 = MS_{within} = 0.00756$. This variability is better expressed as a CV for easy interpretation. The residual variance approximation for $AUC_{0-\infty}$ is 0.00756 ($\hat{\sigma}_W^2 = 0.00756$) and this is equivalent to the within-subject CV of 8.7% assuming that the formulation variances for the within-subject are equal. Similarly, the between-subject CV is 11.5%.

Table 4.20: ANOVA table for the log-transformed C_{max}.

Source of Variation	DF	SS	MS	F statistics	P-value
Between:Subject					
Sequence	1	1.126	1.126	5.17	0.0393
Subject(Sequence)	14	2.8362	0.2026	0.87	0.6021
Within:Within					
Formulation	1	0.0864	0.0864	0.37	0.5525
Period	1	0.0730	0.0730	0.31	0.5847
Residuals	14	3.265	0.233		
Total	31	7.3867			

Table 4.20 gives the ANOVA of the log-transformed C_{max}. The result indicates a significant effect in the sequence (p-value=0.0393) and the subject effect is also not significant (p-value=0.6021) at a 5% level of significance. The non-significant subject effect therefore suggests that there are no inter-subject contrast in clearance and in the C_{max} as well. In addition, there is a non-significant period effect (p-value=0.5847) while the formulation effect is also not significant (p-value=0.5525). This suggests that there are no statistically significant carryover effect, period effect or formulation effects in this bioequivalence study.

The residual variance estimate $\hat{\sigma}_W^2 = MS_{within} = 0.233$. This variability is better expressed as a CV for easy interpretation. The residual variance evaluation for C_{max} is 0.233 ($\hat{\sigma}_W^2 = 0.233$) and this is comparable to the within-subject CV of 48.3% assuming that the formulation variances for the within-subject are equal. Similarly, the between-subject CV is 3.9%.

4.8.1 The Classic Confidence Interval

Table 4.21: The mean of the log-transformed pharmacokinetic parameters.

	Mean	
Parameter	Test	Reference
AUC_{0-t}	4.1441	4.1625
$AUC_{0-\infty}$	4.1995	4.2448
C_{max}	1.1011	0.9972

Table 4.21 presents the mean for the R and the T treatment for each of the pharmacokinetic parameters. Table 4.22 provides the 90% confidence intervals for the different pharmacokinetic parameters. The mean AUC_{0-t} values were 4.1441 and 4.1625 for the T and R drugs, respectively. The mean $AUC_{0-\infty}$ for the T and R drugs were 4.1995 and 4.2448, respectively. The mean C_{max} values were 1.1011 and 0.9972 for the T and R drugs, respectively. The mean AUC_{0-t} and $AUC_{0-\infty}$ for the T drug is lower than the R drug mean for the same parameters. However, the mean C_{max} for the T drug is higher than that of the R drug. The 90% confidence interval for AUC_{0-t} and $AUC_{0-\infty}$ were within the 80-125% limit set by FDA and other regulatory bodies. However, the 90% confidence interval for the C_{max} was not within the 80-125% limit.

Table 4.22: 90% CI for the different pharmacokinetic parameters.

	AUC_{0-t}		$AUC_{0-\infty}$		C_{max}	
	Interval Limits					
	Lower	Upper	Lower	Upper	Lower	Upper
Shortest	98.65	100.46	97.96	99.91	87.49	133.00

Table 4.23: 90% CI and the decision about bioequivalence.

Parameter	Interval Limits		Decision
	Lower	Upper	
AUC_{0-t}	98.65	100.46	ABE claimed
$AUC_{0-\infty}$	97.96	99.91	ABE claimed
C_{max}	87.49	133.00	ABE not claimed

The 90% CI for the log-transformed AUC_{0-t} is given as (98.65,100.46). This CI is within the (80-125) range as per the FDA guidelines. Therefore the R drug and the T drug are declared as average bioequivalent with respect to the AUC_{0-t}.

The associated 90% CI for the $AUC_{0-\infty}$ as indicated in Table 4.23 is (97.96,99.91) and is well within the interval (80-125). The T and R formulations are therefore average bioequivalent for the $AUC_{0-\infty}$.

The associated 90% CI for C_{max} is (87.49,133) which is not contained within the (80-125) range as per the FDA guidelines. ABE cannot be concluded as required by the FDA guidelines as the upper confidence limit is above the required upper limit as per the FDA guidelines. However, for C_{max}, the WHO, the SAMCC and the EMEA adopt a more relaxed equivalence interval of (70-133) for C_{max}. By virtue of this relaxed range used by these organisations, average bioequivalence can be claimed.

4.8.2 Westlake's Symmetric Confidence Interval

For any two formulations to be declared as average bioequivalent using the Westlake symmetric CI, Δ should be less than $0.2\mu_R$ where Δ is defined in Equation 3.36.

Table 4.24: Westlake's CI results for log-transformed AUC_{0-t}, $AUC_{0-\infty}$ and C_{max}.

Parameter	Westlake's Δ value	$0.2\mu_R$
AUC_{0-t}	0.05754786	0.8325
$AUC_{0-\infty}$	0.0866167	0.84896
C_{max}	0.3448097	0.19944

From Table 4.24, the Δ values for the PK parameters AUC_{0-t} and $AUC_{0-\infty}$ are less than $0.2\mu_R$. However, the Δ value for C_{max} is greater than $0.2\mu_R$. Therefore, average bioequivalence is concluded in terms of AUC_{0-t} and $AUC_{0-\infty}$ but not C_{max}.

4.8.3 Schuirmann's TOST Approach

The TOST procedure was introduced by Schuirmann (1987) and was based on Equation 3.13. Two formulations are considered as bioequivalent if H_{01} and H_{02} from Equation 3.13 are both rejected at a significance level which is predetermined.

Table 4.25: The Schuirmann's TOST results for log-transformed PK parameters.

	T_L	T_U	$t(\alpha, n_1 + n_2 - 2)$
AUC_{0-t}	29.08393	-37.82908	1.761
$AUC_{0-\infty}$	26.15258	-36.00039	1.761
C_{max}	1.776692	-0.8514206	1.761

The Schuirmann's TOST results are shown in Table 4.25 for the following PK parameters AUC_{0-t}, $AUC_{0-\infty}$ and C_{max}. From the results, $|T_L|$ and $|T_U|$ for AUC_{0-t} and $AUC_{0-\infty}$ are both greater than $t(0.05, 14) = 1.761$. Since these values are greater than 1.761, the null hypotheses (H_{01} and H_{02}) in Equation 3.13 are both rejected at 5% level of significance. We can conclude bioequiva-

lence according to the ± 20 rule for both AUC_{0-t} and $AUC_{0-\infty}$. However, for C_{max}, $|T_L|$ is greater than 1.761 but $|T_U|$ is less than 1.761. Bioequivalence cannot be claimed using the C_{max}.

4.9   Comparing optimal and regular design

During the sampling times, the general rule of thumb is to collect more blood samples during the absorption phase and peak phase and collect less during the elimination phase. However, the number of samples and the time intervals are generally chosen by the pharmacokineticist and not based on any rule. The regular design of sampling is done without any recognised algorithm and not taking the characteristics of the drug into consideration while the optimal design considers the various characteristics of the drug and uses an algorithm to compute the different times the blood samples are to be collected.

It is very important to verify the quantity of AUC_{0-t} that is missed when the trapezoidal method is used to calculate the AUC_{0-t} using the regular design. The AUC_{0-t} for the regular design and optimal design are presented in Tables 4.26 and 4.27. The second column in Tables 4.26 and 4.27 report the AUC_{0-t} from time 0 to 72 hours for the regular design while column 3 reports the AUC_{0-t} from time 0 to 72 hours for the optimal design. The AUC_{0-t} for the regular design was obtained using the linear interpolation method as explained in Section 2.5 while the AUC_{0-t} for the optimal design was obtained using Equation 2.16 after optimising the sampling times. Column 4 presents the percentage of the unexplained AUC_{0-t} when using the regular design.

Table 4.26: AUC_{0-t} for regular and optimal design.

Subject	AUC_{0-t}(Regular)	AUC_{0-t}(Optimal)	% difference
842	62.4935	67.43587	7.328993
858	60.244	64.62041	6.772489
860	65.0565	63.25278	-2.85161
861	50.333	53.91614	6.645765
862	49.2165	51.74834	4.892601
864	60.549	62.15424	2.582672
869	57.7215	60.51894	4.622421
870	49.312	51.95971	5.095698
842	53.378	65.86338	18.95648
858	63.4214	61.14068	-3.73028
860	78.4096	71.41448	-9.7951
861	56.0545	58.21259	3.707256
862	38.5985	44.36431	12.99651
864	78.99	72.333	-9.20327
869	71.838	63.45344	-13.2137
870	54.3085	63.17753	14.03827
827	56.264	58.80359	4.318767
843	61.1015	60.86652	-0.38606
844	59.2165	61.41677	3.582523
857	60.798	65.92797	7.781174
859	68.9234	67.94529	-1.43956
863	67.7175	68.45833	1.082162
866	66.153	66.2853	0.199592

Table 4.27: AUC_{0-t} for regular and optimal design.

Subject	AUC(Regular)	AUC(Optimal)	% difference
985	58.9715	65.02689	9.312132
827	55.291	58.21236	5.018453
843	73.1955	72.3221	-1.20765
844	72.3501	70.19283	-3.07335
857	66.7527	71.18192	6.222395
859	85.6995	88.02288	2.639518
863	77.544	73.97729	-4.82136
866	59.1085	61.17458	3.377351
985	67.0836	66.62649	-0.68608

The mean percentage of the unexplained AUC_{0-t} while using the regular design is 2.523912 % and a standard deviation of 6.79962 %. The range of the percentage difference is [-13.21372%, 18.95648 %]. The results show that the trapezoidal approximation of the AUC_{0-t} from time 0 to 72 hours for the optimal design is an improvement over the regular design which was used originally. However, there were a few exceptions where the regular design performed better than the optimal design for the subjects. This is an indication that the optimal design could still be improved.

4.10   Summary

This Chapter provided results for the ANOVA, CI, TOST, SA, Wilcoxon-Mann-Whitney test and PopPK model for the blood concentrations which were collected from pigs. The results as discussed in the Chapter indicated that the T drug formulation is average bioequivalent and population bioequivalent to the R drug formulation using the three metrics AUC_{0-t}, $AUC_{0-\infty}$ and C_{max}. The results of the Wilcoxon-Mann-Whitney test on the T_{max} also indicated bioequivalence between the T and R drug formulations.

The second part of the Chapter discussed the development of the PopPK model using a one compartment model with first-order absorption and first-order elimination rate. The model was built using the SAEM algorithm which, according to literature, performs very well. Empirical results and the model diagnostics indicate that the PopPK model is a better fitting model for the data.

The third part of the Chapter discussed the SA algorithm which used the parameter estimates from the PopPK model to obtain optimal sampling times. The empirical results show that the optimal sampling points on average were better than the regular design method of collecting the blood concentrations.

Chapter 5

Conclusion and
Recommendations

5.1 Conclusion

The procedures for assessment of bioequivalence testing have been in existence for the past decades. The methods of bioequivalence assessment have been constantly updated over the years to provide an efficient assessment. These procedures are very important to the pharmaceutical industry as well as the regulatory bodies so that the T drugs could be approved as a replacement for the R drugs which are generally quite expensive. Advances in biopharmaceutical studies made it possible to develop improved drug formulations to enhance efficacy while improving safety and also to create inexpensive drug products which can be marketed once the patent on the R drug product expires. The T drugs have become so important because they are cost effective and help to reduce the medical costs for governments and private consumers.

This dissertation explores several avenues in bioequivalence studies such as the design of the study, different methods of assessment of bioequivalence and where such methods are implemented in drug development. The science of biopharmaceutical studies focuses on differentiating between drug products in order to provide new and improved treatments for the benefit of the public.

Statistical techniques namely the TOST, Westlake's CI, classic CI approach and the Wilcoxon-Mann-Whitney TOST were used to determine ABE between the T and the R drugs. Also, PBE was used to determine if the two drug products were bioequivalent.

In this dissertation, ABE was established between the T and R drug formulations using the AUC since the 90% CI was within the FDA range of acceptance. ABE could not be established under the FDA regulations using the C_{max} since the upper limit of the 90% CI exceeded the FDA CI limits. However, using the regulatory requirements of the EMEA and the SMCC, ABE was established between the T and the R drug formulations since they have a wider CI for the C_{max}. The results of the Wilcoxon-Mann-Whitney TOST using the T_{max} demonstrated ABE between the R and the T drug formulations.

From the hypothesis in Equation 2.9, the null hypothesis is rejected thereby confirming the R and the T drug formulations are population bioequivalent using both the AUC and C_{max}. From the findings in this dissertation, it can be concluded that BE studies play a very important role in health care especially in the developing countries and for those who cannot afford the R drugs. If this concept is embraced and the relevant bodies instituted, the cost of health care will reduce drastically as shown in the literature. The regulatory bodies play a vital role to ensure that even though the T drugs are less expensive and could reduce health care cost as discussed in Chapter 2, T drugs are also as safe and effective as their R drug counterparts in curing their respective illness. This ensures cheap and quality T drugs to be marketed.

However, from the literature, there are limitations to the use of ABE as the sole measure of assessment. It has been stated that it is not enough to prove that two drugs are average bioequivalent but also to use other approaches such as the PBE and IBE to address the limitations of ABE.

5.2 Recommendations

It is recommended that a replicate design study should be used for drugs which are highly variable (intra-subject coefficient of variation greater than 30%) in demonstrating ABE. The use of a replicate design reduces the sample size needed. Also, during the design phase of BE studies, the nature of the drug product should be taken into consideration and should also consider the inter-subject variance homogeneity assumption. The subject-by-formulation interaction variance is assumed to be null during the design stage of the study. However, the variance may not be null. The replicate design offers a benefit that this variance may be separated from intra-subject variation which leads to better understanding of the study outcome.

The pharmacokinetic properties of the drugs should be carefully considered when selecting subjects for the study. Clinical practice should be standardised and conducted in accordance with good practice (time of meals, time of drug administration, etc.).

Also, information on some demographic factors should be provided since they may differ between populations and affect the results. When these assumptions are violated, the power to demonstrate BE would be reduced (Patterson, 2003). BE studies are mostly faced with lack of resources and therefore a small sample of subjects is advised as well as a two-period COD. However, to make a meaningful statistical inference which could lead to a definitive study, a replicate design is preferred which will result in IBE.

Covariate modelling should also be considered as one of the modelling approaches in PopPK modelling, which makes the design of the BE important to include the covariates in the design of the BE study.

References

Aarons, L. (1991). Population pharmacokinetics: theory and practice. *British journal of clinical pharmacology*, Vol.32, No.6, pp.669-670.

Aarts, E. and Korst, J. (1990). Simulated annealing and boltzmann machines. A stochastic approach to combinatorial optimization and neural computing. John Wiley, 272 pages.

Aarts, E., Korst, J. and Michiels, W. (2005). Simulated annealing. Search methodologies, pp.187-210.

Abdel-Rahman, S.M. and Kauffman, R.E. (2004). The integration of pharmacokinetics and pharmacodynamics: understanding dose-response. *Annual Review Pharmacology Toxicology*, Vol. 44, pp.111-136.

Allassonnière, S., Kuhn, E. and Trouvé, A. (2010). Construction of Bayesian deformable models via a stochastic approximation algorithm: a convergence study. *Bernoulli*, Vol. 16, No. 3, pp.641-678.

Altman, D.G. (1999). Statistics in the medical literature: 3. *Statistics in medicine*, Vol. 18, No. 4, pp.487-490.

Aoyagi, N., Ogata, H., Kaniwa, N. and Ejima, A. (1984). Bioavailability of griseo- fulvin plain tablets in stomach-emptying controlled rabbits and the correlation with bioavailability in humans. *Journal of pharmacobio-dynamics* , Vol. 7, No. 9, pp.630-640.

Asyali, M.H. (2010). Design of optimal sampling times for pharmacokinetic trials via spline approximation. *Turkish Journal of Electrical Engineering and Computer Science*, Vol. 18, No. 6, pp.1019–1030.

Bauer, L. (2008). Applied clinical pharmacokinetics. Mcgrawhill, 840 pages.

Bauer, R.J., Guzy, S. and Ng, C. (2007). A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. *The AAPS journal*, Vol. 9, No. 1, pp.E60–E83.

Benet, L.Z. and Goyan, J.E. (1995). Bioequivalence and narrow therapeutic index drugs. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, Vol. 15, No. 4, pp.433-440.

Bergstrand, M., Hooker, A.C., Wallin, J.E. and Karlsson, M.O. (2011). Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models. *The AAPS journal*, Vol.13, No.2, pp.143–151.

Bois, F.Y., Tozer, T.N., Hauck, W.W., Chen, M.L., Patnaik, R. and Williams, R.L. (1994). Bioequivalence: performance of several measures of extent of absorption. *Pharmaceutical research*, Vol.11, No.5, pp.715–722.

Bonate, P.L. and Steimer, J.L. (2006). Pharmacokinetic-pharmacodynamic modelling and simulation. New York: Springer, pp 205-265.

Buehler, G. and Director, A.D. (2010). History of bioequivalence for critical dose drugs, *US Food and Drug Administrations hemsida [www]. Hämtat från: http://www. fda. gov [uppdaterad 2010-04-13, citerad 2014-03-15].*

Buice, R.G., Subramanian, V.S., Duchin, K.L. and Uko-Nne, S. (1996). Bioequivalence of a highly variable drug: an experience with nadolol. *Pharmaceutical research*, Vol.13, No.7, pp.1109–1115.

Byron, J. and Kenward, M.G. (2003). Design and analysis of crossover trials, 408 pages.

Cabana, B.E. (1983). Assessment of 75/75 rule: FDA viewpoint. *Journal of pharmaceutical sciences*, Vol.72, No.1, pp.98–99.

Caccia, S. and Garattini, S. (1990). Formation of active metabolites of psychotropic drugs. *Clinical pharmacokinetics*, Vol. 18, No. 6, pp.434-459.

Calvert, R.T. (1996). Bioequivalence and generic prescribing: A pharmacy view. *Journal of pharmacy and pharmacology*, Vol.48, No.1, pp.9–10.

Cartwright, A.C. (1991). International harmonization and consensus DIA meeting on bioavailability and bioequivalence testing requirements and standards. *Drug Information Journal*, Vol.25, No.4, pp.471–482.

Chen, M.L. and Pelsor, F.R. (1991). Half-life revisited: Implication in clinical trials and bioavailability/bioequivalence evaluation. *Journal of pharmaceutical sciences*, Vol. 80, No. 4, pp.406-408.

Choi, L., Caffo, B. and Rohde, C. (2007). Optimal sampling times in bioequivalence studies using a simulated annealing algorithm. *Statistics and Computing*, Vol. 17, No. 4, pp.337-347.

Chow, S.C. and Liu, J.P. (2008). Design and analysis of bioavailability and bioequivalence studies. CRC Press, 760 pages.

Chow, S.C. and Tse, S.K. (1990). Outlier detection in bioavailability/bioequivalence studies. *Statistics in medicine*, Vol.9, No.5, pp.549–558.

Chow, S.C. and Tse, S.K. (1991). On variance estimation in assay validation. *Statistics in Medicine*, Vol.10, pp.1543–1553.

Comets, E., Brendel, K. and Mentré, F. (2010). Model evaluation in nonlinear mixed effect models, with applications to pharmacokinetics. *Journal de la Société Française de Statistique*, Vol.151, No.1, pp.106–128.

Cook, R.D. (1977). Detection of influential observation in linear regressions. *Technometrics*, Vol.19, No.1, pp.15–18.

Cook, R.D. and Weisberg, S. (1982). Residuals and influence in regression. New York: Chapman and Hall, 230 pages.

Cornell, R.G. (1990). The evaluation of bioequivalence using nonparametric procedures. *Communications in Statistics-Theory and Methods*, Vol. 19, No. 11, pp.4153–4165.

D'Argenio, D.Z. (1981). Optimal sampling times for pharmacokinetic experiments. *Journal of pharmacokinetics and biopharmaceutics*, Vol.9, No.6, pp.739–756.

Delyon, B., Lavielle, M. and Moulines, E. (1999). Convergence of a stochastic approximation version of the EM algorithm. *Annals of statistics*, pp.94–128.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, pp.1–38.

Dixon, W.J. (1953). Processing data for outliers. *Biometrics*, Vol.9, No.1, pp.74–89.

Draper, N.R. and John, J.A. (1981). Influential observations and outliers in regression. *Technometrics*, Vol.23, No.1, pp.21–26.

El-Tahtawy, A.A., Tozer, T.N., Harrison, F., Lesko, L. and Williams, R. (1998). Evaluation of bioequivalence of highly variable drugs using clinical trial simulations. II: Comparison of single and multiple-dose trials using AUC and Cmax. *Pharmaceutical research*, Vol.15, No.1, pp.98–104.

EMEA (2001). Guidance for the conduct of bioequivalence studies for veterinary medicinal products. pp. 1-11

Endrenyi, L. and Midha, K.K. (1998). Individual bioequivalence—has its time come?. *European journal of pharmaceutical sciences*, Vol. 6, No. 4, pp.271–277.

Fan, J. and de Lannoy, I.A. (2014). Pharmacokinetics, *Biochemical Pharmacology*. Vol. 87, No. 1, pp.93–120.

F.D.A. (1992). Guidance for industry: Statistical procedures for bioequivalence studies using a standard two treatment crossover design, US Food and Drug Administration, Center for Drug Evaluation and Research (CDER).

F D.A. (1997). Guidance for industry: Statistical approaches to establishing bioequivalence, Center for Drug Evaluation and Research (CDER), United States Food and Drug Administration.

F.D.A. (1999). Guidance for industry: Population pharmacokinetics, US Department of health and Human Services, US Food and Drug Administration, Center for Drug Evaluation and Research (CDER).

F. D.A. (2001). Guidance for industry: In vivo bioequivalence studies based on population and individual bioequivalence approaches, US Food and Drug Administration, Center for Drug Evaluation and Research (CDER).

F. D.A. (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.

Ghosh, P. and Rosner, G. L. (2007). A semi-parametric bayesian approach to average bioequivalence. *Statistics in medicine*, Vol. 26, No. 6, pp.1224-1236.

Gibaldi, M. and Perrier, D. (1982). Multicompartment models. *Pharmacokinetics*, Vol. 2, pp.45–111.

Gray, A.L. (2009). Medicine pricing interventions–the South African experience. *Southern Med Review*, Vol.2, No.2, pp.15–19.

Hauck, W.W. and Anderson, S. (1984). A new statistical procedure for testing equivalence in two-group comparative bioavailability trials. *Journal of Pharmacokinetics and Pharmacodynamics*, Vol.12, No.1, pp.83–91.

Hauck, W.W. and Anderson, S. (1992). Types of bioequivalence and related statistical considerations. *International journal of clinical pharmacology, therapy, and toxicology*, Vol.30, No.5, pp.181–187.

Hauschke, D., Steinijans, V. and Pigeot, I. (2007). Bioequivalence studies in drug development: methods and applications. John Wiley & Sons, 321 pages.

Haynes, J.D. (1981). Statistical simulation study of new proposed uniformity requirement for bioequivalency studies. *Journal of pharmaceutical sciences*, Vol. 70, No. 6, pp.673–675.

Henry, D. and Lexchin, J. (2002). The pharmaceutical industry as a medicines provider. *The Lancet*, Vol.360, No.9345, pp.1590–1595.

Herchuelz, A. (1996). Bioequivalence assessment and the conduct of bioequivalence trials: a European point of view. *European journal of drug metabolism and pharmacokinetics*, Vol.21, No.2, pp.149–152.

Jones, B. and Kenward, M.G. (2014). Design and analysis of crossover trials. CRC Press, 438 pages.

Jones, B., Wang, J., Jarvis, P. and Byrom, W. (1999). Design of crossover trials for pharmacokinetic studies. *Journal of statistical planning and inference*, Vol. 78, No. 1, pp. 307-316.

Kamerow, D. (2011). The pros and cons of generic drugs. *BMJ: British Medical Journal (Online)*, page 343.

Karim, S.A., Pillai, G., Ziqubu-Page, T.T., Cassimjee, M.H. and Morar, M.S. (1996). Potential savings from generic prescribing and generic substitution in South Africa. *Health policy and planning*, Vol.11, No.2, pp.198–202.

Karlsson, M.O. and Holford, N.I.C.K. (2008). June. A tutorial on visual predictive checks. *In Annual Meeting of the Population Approach Group in Europe*, Vol.17.

Katz, D. and D'Argenio, D.Z. (1983). Experimental design for estimating integrals by numerical quadrature, with applications to pharmacokinetic studies. *Biometrics*, Vol.39, No.3, pp.621–628.

Kimura, T. and Higaki, K. (2002). Gastrointestinal transit and drug absorption. *Biological and Pharmaceutical Bulletin*, Vol.25, No.2, pp.149–164.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983). Optimisation by simulated annealing. *Science*, Vol.220, No.4598, pp.671–680.

Kirkwood, B. and Sterne, J. (2003). Essential medical statistics. John Wiley and Sons, 501 pages.

Kirkwood, T.B. and Westlake, W. (1981). Bioequivalence testing–a need to rethink. *Biometrics*, Vol. 37, No. 3, pp.589–594.

Koch, G.G. (1972). The use of non-parametric methods in the statistical analysis of the two-period change-over design. *Biometrics*, pp. 577–584.

Kong, F.H. and Gonin, R. (2000). Optimal sampling times in bioequivalence tests. *Journal of biopharmaceutical statistics*, Vol.10, No.1, pp.31–44.

Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis*, Vol.49, No.4, pp.1020–1038.

Lacey, L.F., Bye, A. and Keene, O.N. (1995). Glaxo's experience of different absorption rate metrics of immediate release and extended release dosage forms. *Drug information journal*, Vol.29, No.3, pp.821-840.

Lavielle, M. (2014). Mixed effects models for the population approach: models, tasks, methods and tools. CRC, press 383 pages.

Lehmann, E.L. and Casella, G. (2006). Theory of point estimation. Springer Science and Business Media, 590 pages.

Levy, N. (1986). Bioequivalence of solid oral dosage forms. A presentation to the US Food and Drug Administration Hearing on Bioequivalence of Solid Oral Dosage Forms. September, pp.9–11.

Liu, J.P. and Weng, C.S. (1991). Detection of outlying data in bioavailability/bioequivalence studies. *Statistics in Medicine*, Vol.10, No.9, pp.1375-1389.

Lockyer, M., Al-Dgither, S., Al-Gaai, E., Yousuf, A. and Hammami, M.M. (2005). Bioequivalence of two oral formulations of gatifloxacin tablets in healthy male volunteers. *J Appl Res*, Vol.5, No.1, pp.73–9.

Martinez, M.N. and Jackson, A.J. (1991). Suitability of various noninfinity area under the plasma concentration–time curve (auc) estimates for use in bioequivalence determinations: Relationship to AUC from zero to time infinity (AUC0–inf). *Pharmaceutical research*, Vol. 8, No. 4, pp.512–517.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, Vol.21, No.6, pp.1087–1092.

Metzler, C. and Huang, D. (1983). Statistical methods for bioavailability and bioequivalence. *Clinical Research and Regulatory Affairs*, Vol.1, No.2, pp.109–132.

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

Midha, K.K. and McKay, G. (2009). Bioequivalence; its history, practice, and future. *The AAPS journal*, Vol.11, No.4, pp.664–670.

Midha, K.K., Rawson, M.J. and Hubbard, J.W. (1997). Individual and average bioequivalence of highly variable drugs and drug products. *Journal of pharmaceutical sciences*, Vol.86, No.11, pp.1193-1197.

Midha, K.K., Ormsby, E.D., Hubbard, J.W., McKay, G., Hawes, E.M., Gavalas, L. and McGilveray, I.J. (1993). Logarithmic transformation in bioe-

quivalence: application with two formulations of perphenazine. *Journal of pharmaceutical sciences*, Vol.82, No.2, pp.138-144.

Owen, D.B. (1965). The power of Student's t-test. *Journal of the American Statistical Association*, Vol.60, No.309, pp.320–333.

Pabst, G. and Jaeger, H. (1990). Review of methods and criteria for the evaluation of bioequivalence studies. *European journal of clinical pharmacology*, Vol.38, No.1, pp.5–10.

Papadimitriou, C.H. and Steiglitz, K. (1998). Combinatorial optimization: algorithms and complexity. Courier Corporation.

Patterson, S.D. (2001). A Review Of The Development Of Biostatistical Design And Analysis Techniques For Assessing In Vivo Bioequivalence: Part One. *Indian Journal of Pharmaceutical Sciences*, Vol.63, No.2, pp.81-100.

Patterson, S.D. (2003). Statistical aspects of bioequivalence assessment in the pharmaceutical industry (Doctoral dissertation, University of London).

Peter, J.B. and Kjell, A.D.(2001). Mathematical statistics: basic ideas and selected topics. Vol.1, Second edition.

Pham, D. and Karaboga, D. (2012). Intelligent optimisation techniques: genetic algorithms, tabu search, simulated annealing and neural networks. Springer Science and Business Media.

Phillips, K.F. (1990). Power of the two one-sided tests procedure in bioequivalence. *Journal of pharmacokinetics and biopharmaceutics*, Vol.18, No.2, pp.137-144.

Pocock, S.J. (2013). Clinical trials: a practical approach. John Wiley and Sons, 278 pages.

Purich, E. (1980). Bioavailability/bioequivalence regulations: An FDA perspective. Drug Absorption and Disposition: Statistical considerations. *Ameri-*

can Pharmaceutical Association, Academy of Pharmaceutical Sciences, Washington, DC, pp.115–137.

Rani, S. and Pargal, A. (2004). Bioequivalence: An overview of statistical concepts.

Rebull, J.O., Sanchez, M.P., Pla, A.S. and Pla, J.L.C. (2008). On equivalence and bioequivalence testing. *Statistics and operations research transactions*, Vol.32, No.2, pp.151–176.

Rescigno, A. and Powers, J.D. (1998). AUC and Cmax are not sufficient to prove bioequivalence. *Pharmacological research*, Vol.37, No.2, pp.93–95.

Rheinstein, P.H. (1990). Therapeutic inequivalence. *Drug safety*, Vol.5, No.1, pp.114-119.

Rodda, B.E. (1986). Bioequivalence of solid oral dosage forms. *A Presentation to US Food and Drug Administration Hearing on Bioequivalence of Solid Oral Dosage Forms, Pharmaceutical Manufacturers Association*, September, pp.12–5.

Rowland, M. (1980). Intra-individual variability in pharmacokinetics, Towards Better Safety of Drugs and Pharmaceutical Products. *Elsevier*, Amsterdam, pp.143–151.

Rowland, M. and Tozer, T. (1995). Clinical pharmacokinetics: Concepts and applications. Williams and Wilkins, 839 pages.

Saul, S. (2007). More generics slow rise in drug prices. *The New York Times*, 8 August.

Saul, S. and Berenson, A. (2007). Maker of Lipitor digs in to fight generic rival. *New York Times*, 3 November.

Schuirmann, D.L. (1981). On hypothesis-testing to determine if the mean of a normal-distribution is contained in a known interval. *Biometrics*, Vol.37, No.3, pp. 617-617.

Schuirmann, D.J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of pharmacokinetics and biopharmaceutics*, Vol.15, No.6, pp.657–680.

Senn, S.S. (2001). Crossover trials in drug development: theory and practice. *Journal of statistical planning and inference*, Vol.96, No.1, pp.29–40.

Senn, S.S. (2008). Statistical issues in drug development, Vol.69. John Wiley and Sons, 524 pages.

Shargel, L., Wu-Pong, S. and Yu, B.C.A. (2004). Applied biopharmaceutics and pharmacokinetics. McGraw-Hill, New York, pp.412-450.

Sheiner, L.B. and Beal, S.L. (1985). Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. *Journal of pharmacokinetics and biopharmaceutics* , Vol.13, No.2, pp.185–201.

Sheiner, L.B., Rosenberg, B. and Marathe, V.V. (1977). Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *Journal of pharmacokinetics and biopharmaceutics*, Vol.5, No.5, pp.445–479.

Sheiner, L.B., Rosenberg, B. and Melmon, K.L. (1972). Modelling of individual pharmacokinetics for computer-aided drug dosage. *Computers and Biomedical Research*, Vol.5, No.5, pp.441–459.

Sheiner, L.B. and Steimer, J.L. (2000). Pharmacokinetic/pharmacodynamic modeling in drug development. *Annual review of pharmacology and toxicology*, Vol.40, No.1, pp.67–95.

Steinijans, V.W., Sauter, R., Hauschke, D. and Elze, M. (1995). Metrics to characterize concentration-time profiles in single-and multiple-dose bioequivalence studies. *Drug information journal*, Vol.29, No.3, pp.981–987.

Strom, B.L. (1987). Generic drug substitution revisited. *New England Journal of Medicine*, Vol.316, No.23, pp.1456–1462.

Sun, X. (2010). Diagnostics for nonlinear models with application to population pharmacokinetic modeling.

Teorell, T. (1937). Kinetics of distribution of substances administered to the body, I: the extravascular modes of administration. *Archives internationales de pharmacodynamie et de therapie*, Vol.57, pp.205–225.

Terry, S., Gould, J., McManus, J. and Prescott, L. (1982). Absorption of penicillin and paracetamol after small intestinal bypass surgery. *European journal of clinical pharmacology*, Vol.23, No.3, pp.245–248.

Tokdar, S.T. and Kass, R.E. (2010). Importance sampling: a review. Wiley Interdisciplinary Reviews. *Computational statistics*, Vol.2, No.1, pp.54–60.

Tsang, Y.C., Pop, R., Gordon, P., Hems, J. and Spino, M. (1996). High variability in drug pharmacokinetics complicates determination of bioequivalence: experience with verapamil. *Pharmaceutical research*, Vol.13, No.6, pp.846–850.

Wang, D. and Bakhai, A. (2006). Clinical trials: a practical guide to design, analysis, and reporting. Remedica, London, 498 pages.

Wellek, S. (1993). Basing the analysis of comparative bioavailability trials on an individualized statistical definition of equivalence. *Biometrical journal*, Vol.35, No.1, pp.47-55.

Welling, P. (1984). Effects of gastrointestinal disease on drug absorption. *Pharmadynamic Basis for Drug Treatment*, pp.29-47.

Welling, P. (1997). Pharmacokinetics: processes, mathematics, and applications. Amer Chemical Society.

Welty, T.E., Pickering, P.R., Hale, B.C. and Arazi, R. (1992). Loss of seizure control associated with generic substitution of carbamazepine. *Annals of Pharmacotherapy*, Vol.26, No.6, pp.775-777.

Westlake, W.J. (1972). Use of confidence intervals in analysis of comparative bioavailability trials. *Journal of Pharmaceutical Sciences*, Vol.61, No.8, pp.1340–1341.

Westlake, W.J. (1976). Symmetrical confidence intervals for bioequivalence trials. *Biometrics*, pp.741–744.

Westlake, W.J. (1979). Statistical aspects of comparative bioavailability trials. *Biometrics*, pp.273–280.

Westlake, W.J. (1981). Bioequivalence testing–a need to rethink. *Biometrics*, Vol.37, No.3, pp.589–594.

Wilkins, J., Karlsson, M. and Jonsson, E.N. (2006). Patterns and power for the visual predictive check. In 15th meeting of the Population Approach Group in Europe, Brugges, Belgium, Vol.1029.

Wu, C.J. (1983). On the convergence properties of the EM algorithm. *The Annals of statistics*, pp.95–103.

Appendix A

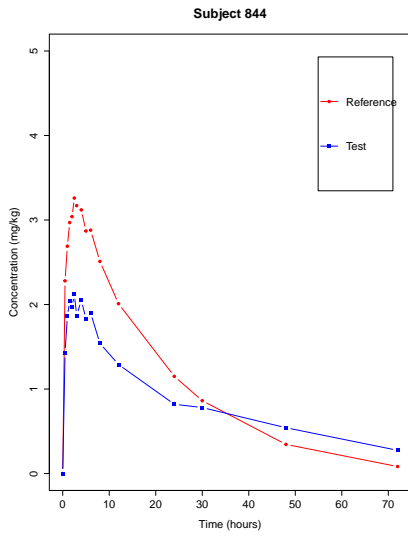# The CTC for the subjects and data for calculating $AUC_{0-\infty}$
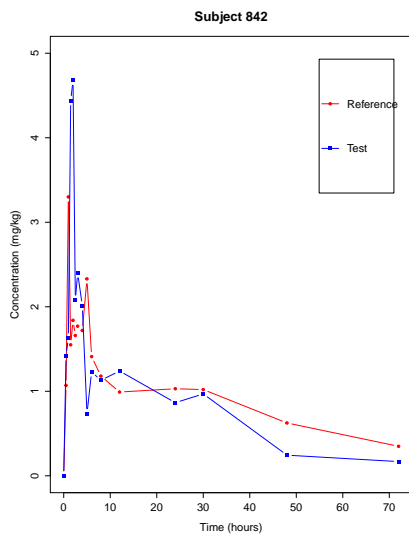
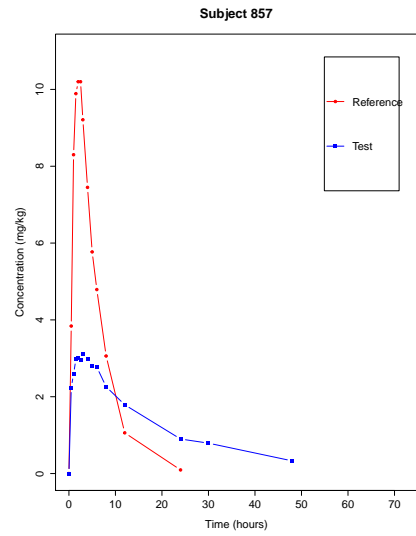Figure A.1: Concentration profile for T and R formulation in sequence 1
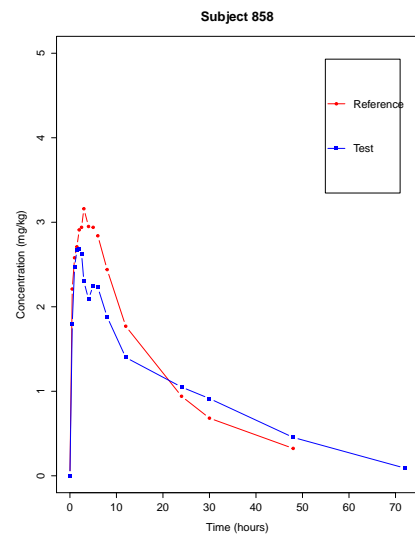
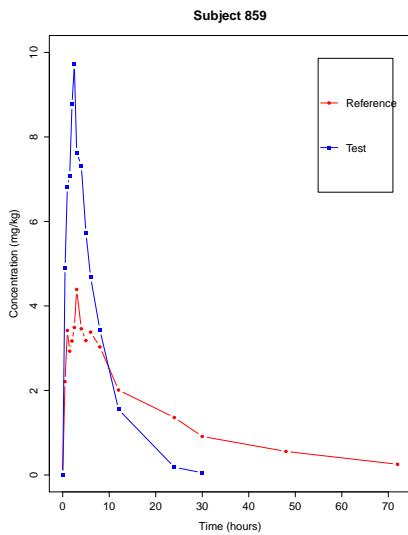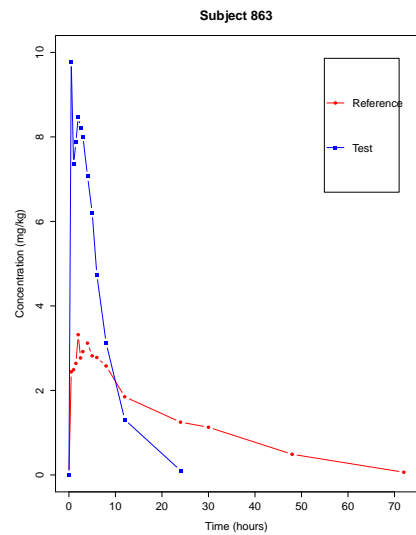Figure A.2: Concentration profile for T and R formulation in sequence 1
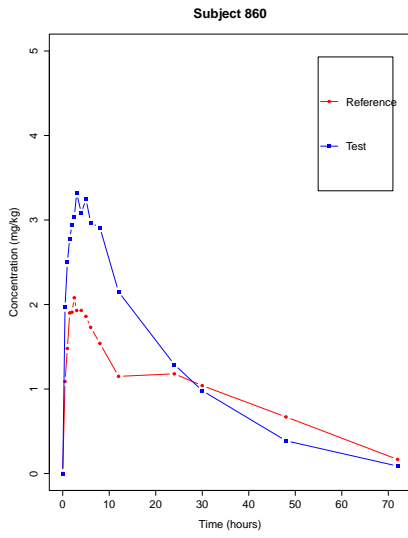
Figure A.3: Concentration profile for T and R formulation in sequence 1
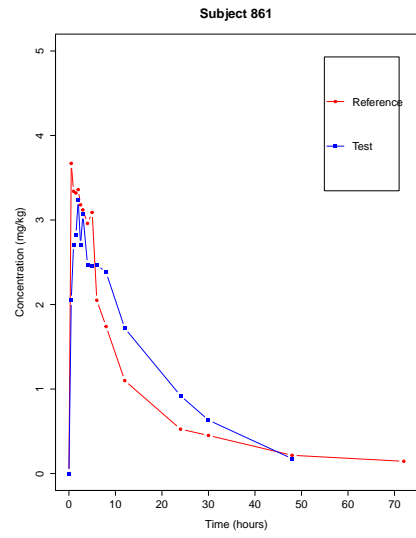
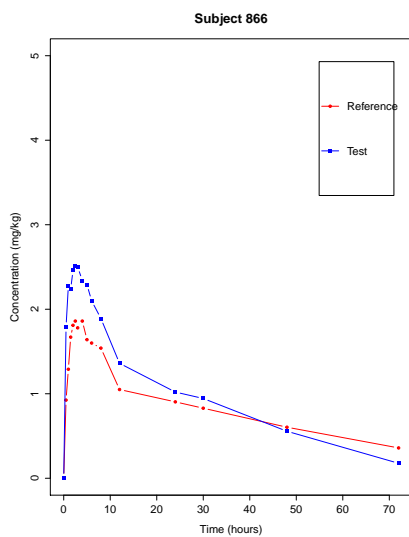Figure A.4: Concentration profile for T and R formulation in sequence 1
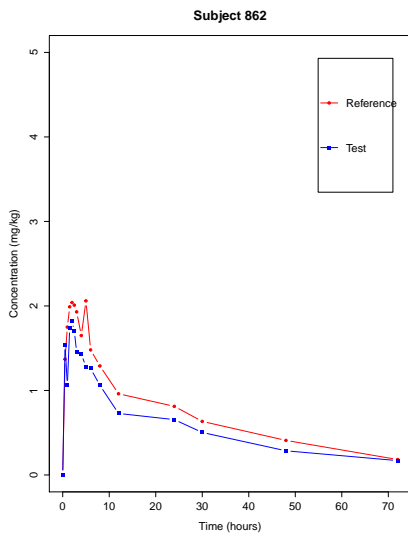
Figure A.5: Concentration profile for T and R formulation in sequence 1
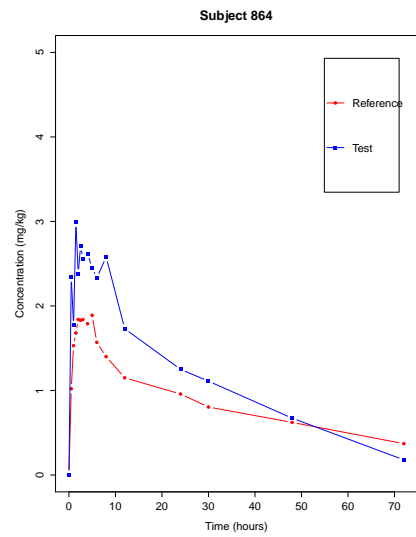
Figure A.6: Concentration profile for T and R formulation in sequence 1
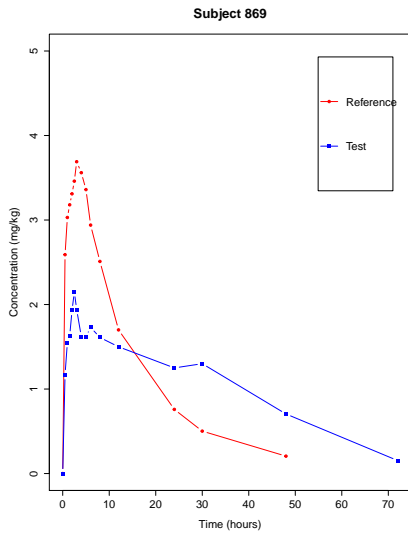
Figure A.7: Concentration profile for T and R formulation in sequence 1
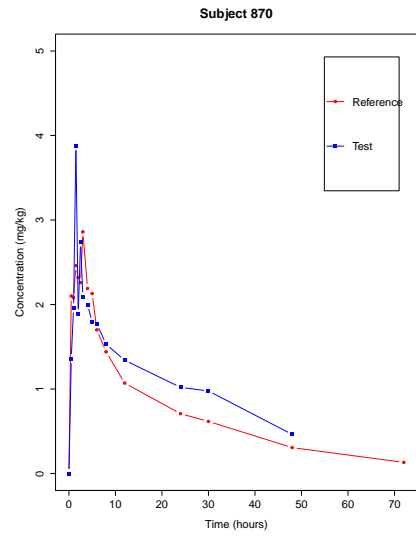
Figure A.8: Concentration profile for T and R formulation in sequence 1
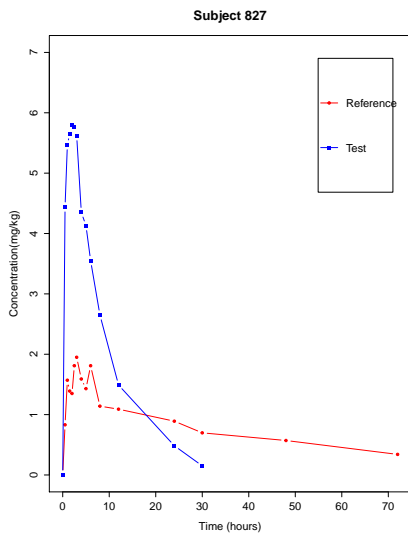
Figure A.9: Concentration profile for T and R formulation in sequence 2
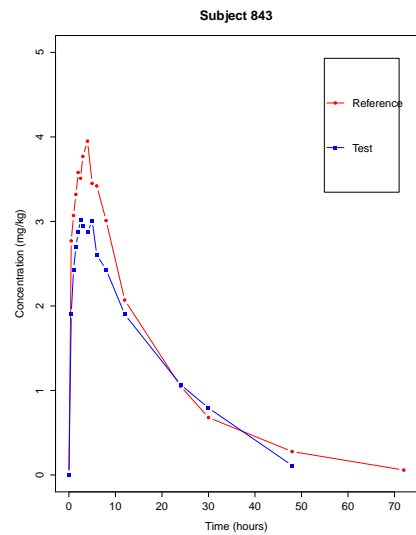
Figure A.10: Concentration profile for T and R formulation in sequence 2
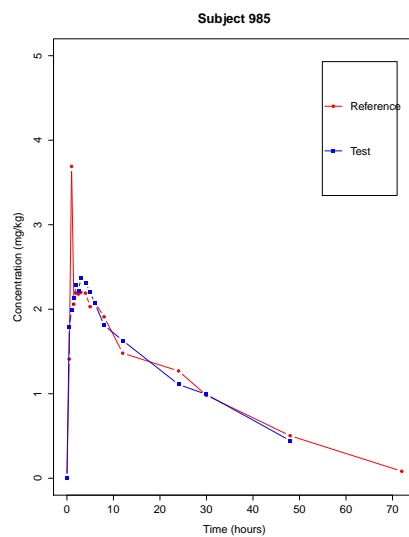
162

Figure A.11: Concentration profile for T and R formulation in sequence 2

Figure A.12: Concentration profile for T and R formulation in sequence 2

Figure A.13: Concentration profile for T and R formulation in sequence 2

Figure A.14: Concentration profile for T and R formulation in sequence 2

Figure A.15: Concentration profile for T and R formulation in sequence 2

Figure A.16: Concentration profile for T and R formulation in sequence 2

Table A.1: Data for calculating $AUC_{0-\infty}$

Subject number	Interval for k_e estimation (λ)	λ	$t_{1/2}$	C_t	AUC_{0-t}	$AUC_{0-\infty}$
842	6-72	0.0188846	36.70	0.349	62.4935	80.9742
858	4 − 48	0.05242	13.22	0.324	60.244	66.4248
860	5 − 72	0.031445	22.04	0.167	65.0565	70.36736
861	6-72	0.040367	17.17	0.145	50.333	53.92504
862	6 − 72	0.029845	23.22	0.183	49.2165	55.34818
864	6 − 72	0.020561	33.71	0.371	60.549	78.59287
869	6 − 48	0.063876	10.85	0.206	57.7215	60.94650
870	6 − 72	0.037387	18.85	0.132	49.312	52.84264
827	8 − 72	0.01883	36.81	0.341	55.291	73.4004
843	5-72	0.06070	11.42	0.058	73.1955	74.1510
844	5-72	0.05239	13.23	0.0828	72.3501	73.9506
857	3-24	0.219897	3.15	0.0942	66.7527	67.1811
859	4-72	0.039262	17.65	0.253	85.6995	92.1434
863	6-72	0.053395	12.98	0.0645	77.544	78.7520
866	5-72	0.021901	31.65	0.359	59.1085	75.5004
985	6-72	0.04552	15.23	0.0828	67.0836	68.9026

Table A.2: Data for calculating $AUC_{0-\infty}$

Subject Number	Interval	lambda	half-life	C	AuC	AUC infity
842	4-72	0.031385	22.09	0.168	53.378	58.7309
858	6-72	0.04476	15.49	0.0912	63.4214	65.4589
860	6-72	0.052737	13.14	0.0883	78.4096	80.08395
861	8-72	0.063779	10.87	0.176	56.0545	58.8140
862	6-72	0.029050	23.86	0.170	38.5983	44.4499
864	12-72	0.037320	18.57	0.179	78.99	83.78636
869	8-72	0.03518	19.70	0.154	71.838	76.21549
870	4-48	0.030158	22.98	0.463	54.3085	69.66098
827	4-30	0.12395	5.59	0.151	56.264	57.4822
843	6-48	0.071681	9.67	0.112	61.1015	62.6640
844	8-72	0.02562	27.05	0.275	59.2165	69.9503
857	4-48	0.05036	13.76	0.330	60.798	67.3508
859	3-30	0.18584	3.73	0.0513	68.9234	69.1994
863	3-24	0.21020	3.30	0.105	67.7175	68.2170
866	6-72	0.034855	19.89	0.177	66.153	71.2319
985	4-48	0.03546	19.55	0.445	59.9715	72.5209

166

Appendix B

Diagnostic plots from population Pharmacokinetic Modelling
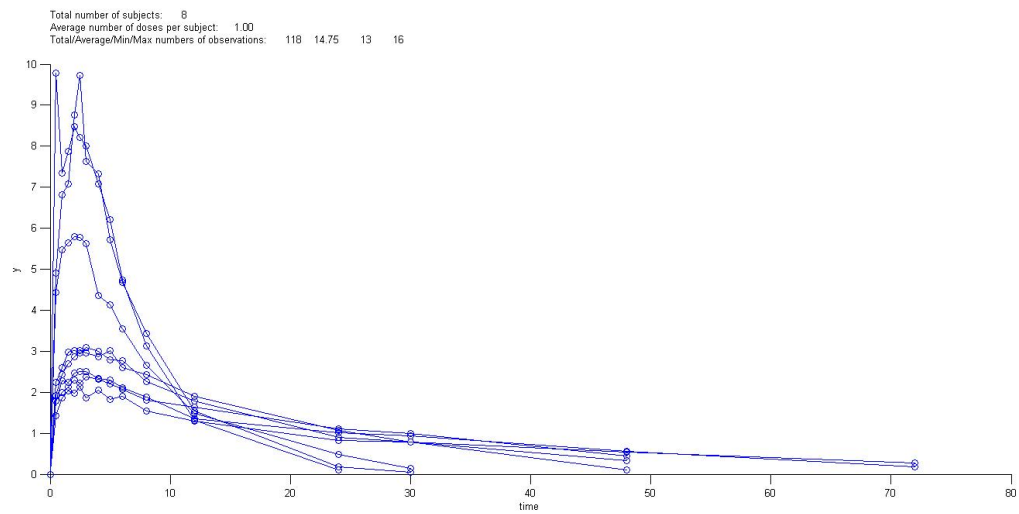
Figure B.1: Spaghetti plot with parameters Ka, V and Cl
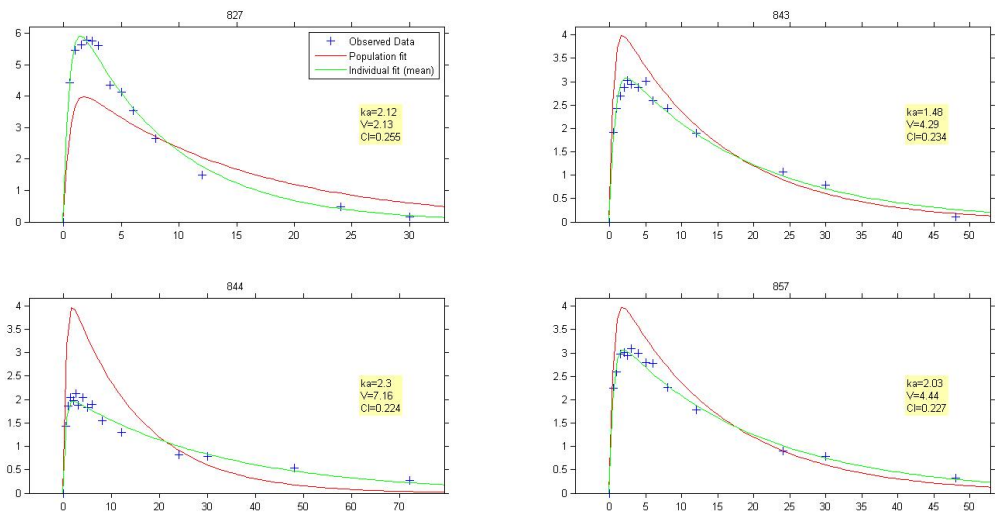
Figure B.2: Individual and population fits for subjects
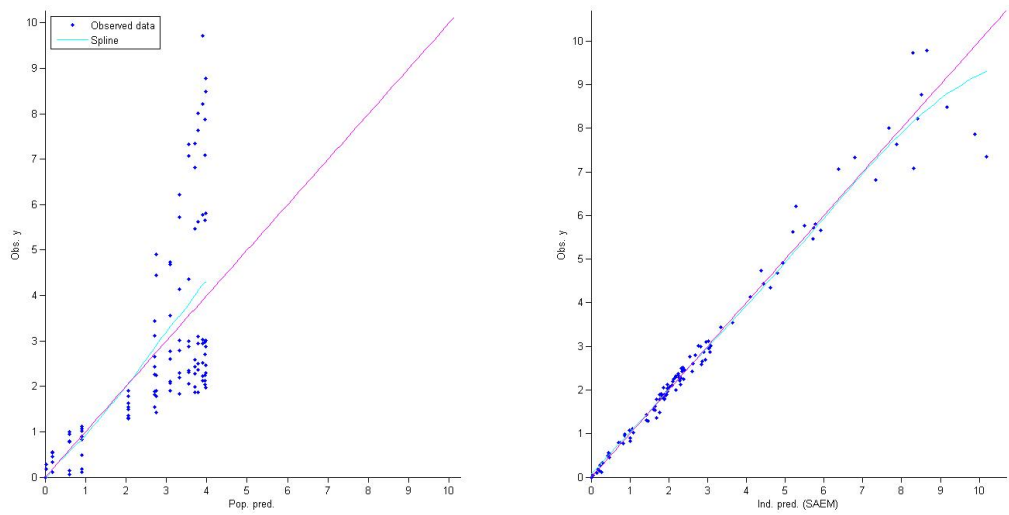

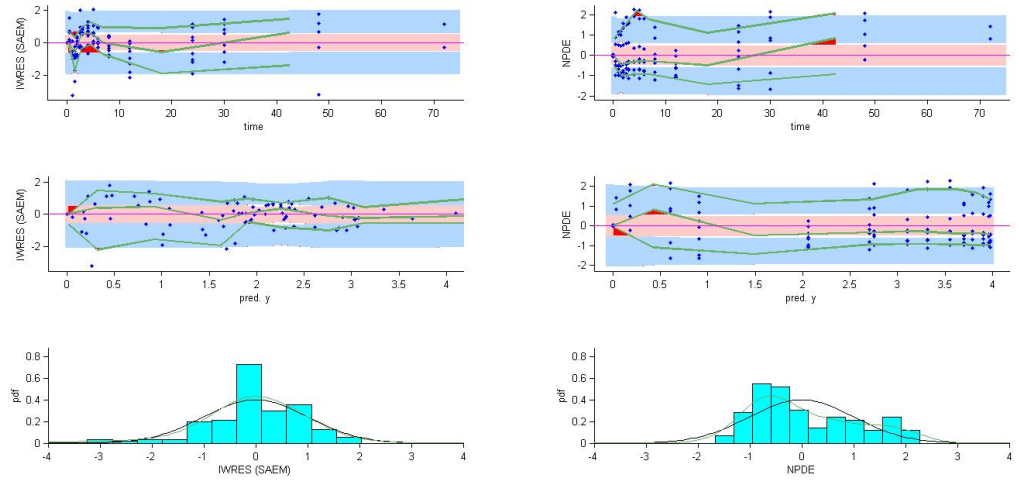Figure B.3: Observation versus predictions
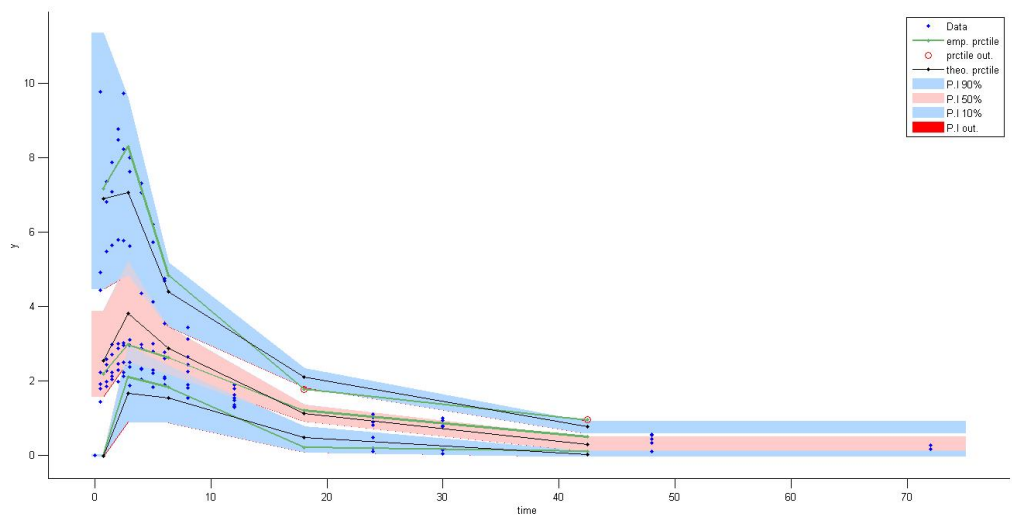
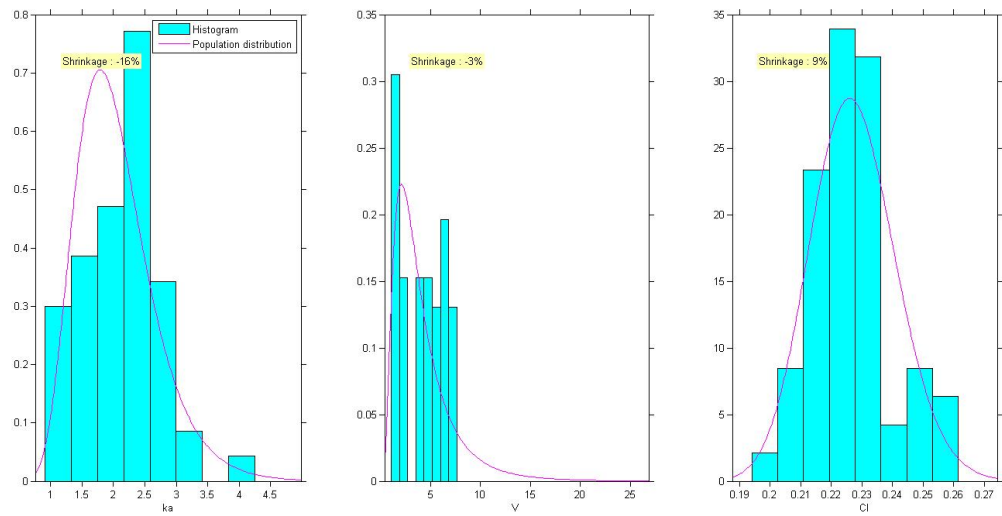Figure B.4: Individual weighted residuals and NPDE

Figure B.5: Visual predictive check

Figure B.6: Population distribution of the estimated individual parameters and the empirical distributions of the individual parameters simulated with conditional distributions
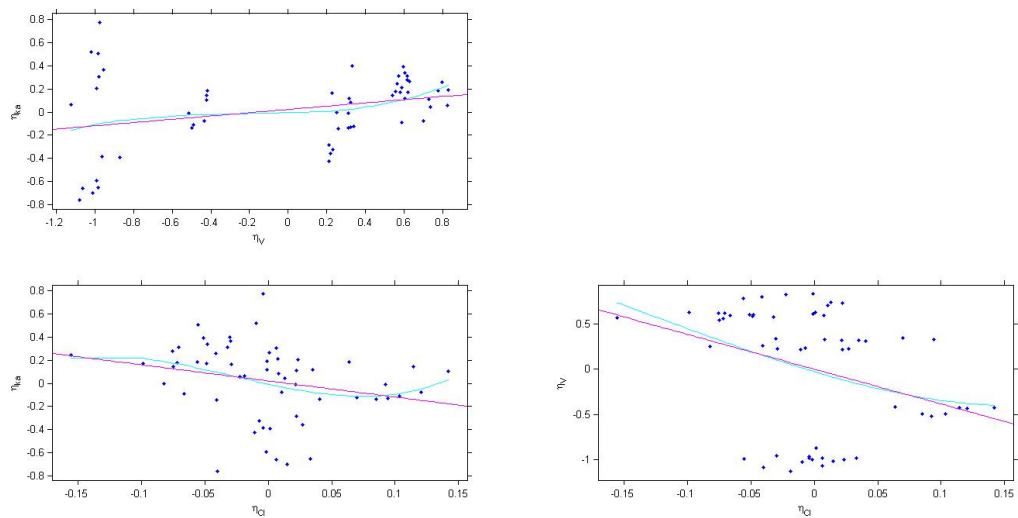

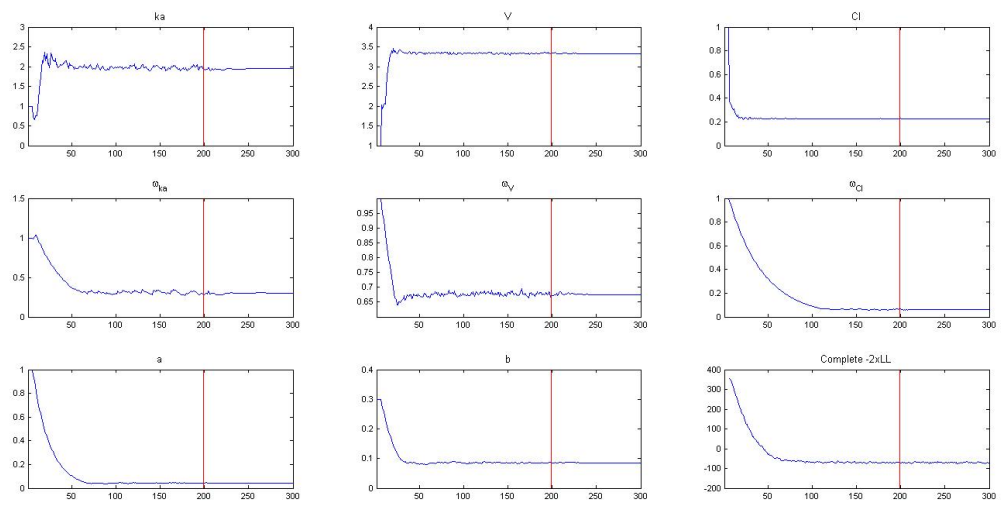Figure B.7: Joint distribution of the random effects

Figure B.8: SAEM convergence

Appendix C

R and MATLAB codes

R code for ABE and PBE

chow$Sequence<−as.factor(chow$Sequence)

chow$Period<−as.factor(chow$Period)

chow$Subject<−as.factor(chow$Subject)

```
###converting AUC and Conc to natural log in chow——#
for( i in 1:nrow(chow)){
chow$lnAUCT[i]<−(log(chow$AUCT[i]))
}

for( i in 1:nrow(chow)){
chow$lnAUCINF[i]<−(log(chow$AUCINF[i]))
}

for( i in 1:nrow(chow)){
chow$lnCMAX[i]<−(log(chow$CMAX[i]))
}

clc
clear all
```

```matlab
format long g
t=[0 0.5 1 1.5 2 2.5 3 4 5 6 8 12 24 24 30 48 72];
ka=0.0393; k=2.71; D=15; V=0.0868;F=1;
c=ka*F*D/(V*(ka-k));
f=@(t) c.*(exp(-k.*t)-exp(-ka.*t));
T=20;
[a,b]=size(t);
m=b-2;
for i=2:24
omega=f(t);
T=T/log(i);
k1=randi([2,m],1);
tnew1=[t(k1-1), t(k1+1)];
r=unifrnd(t(k1-1), t(k1+1));
tnew=t;
tnew(k1)=r;
sigma=f(tnew);
p=min([exp(-(sigma-omega)./T),1]);
u=rand;
if u<=p
t=tnew;
else
end
end
display(t')
plot(t,f(t))
xlabel('time(hours)')
ylabel('concentration(mg/kg)')
grid on

clc
```

```
clear all
format long g
ka=0.0393; k=2.71; D=15; V=0.0868;F=1;
c=ka*F*D/(V*(ka-k));
f=@(t) c.*(exp(-k.*t)-exp(-ka.*t));

proc glm
data=WORK.DATA1;
class Formulation Period Sequence Subject;
model lnAUC= Sequence Subject(Sequence) Period Formulation;
random Subject/ test;
lsmeans Formulation/pdiff cl alpha=0.1;
run;


proc mixed
data=WORK.DATA1;
class Subject Sequence Period Formulation;
model lnAUC=Sequence Period Formulation;
random Subject(Sequence);
lsmeans Formulation/pdiff cl alpha=0.1;
estimate 'T/R' Formulation 1 2 / cl alpha=0.1;
run;

setwd("C:\\Users\\Administrator\\Desktop\\optimisation")
chow= read.csv("C:\\Users\\Administrator\\Desktop\\optimisation\\chow

chow$Sequence<-as.factor(chow$Sequence)
chow$Period<-as.factor(chow$Period)
chow$Subject<-as.factor(chow$Subject)

###converting AUC and Conc to natural log in chow——#
```

```
for ( i in 1:nrow(chow)){
chow$lnAUCT[i]<-(log(chow$AUCT[i]))
}


for ( i in 1:nrow(chow)){
chow$lnAUCINF[i]<-(log(chow$AUCINF[i]))
}



for ( i in 1:nrow(chow)){
chow$lnCMAX[i]<-(log(chow$CMAX[i]))
}


tab.n<-aggregate(chow$lnAUCT, list(Sequence=chow$Sequence, Period=chow$
n1<-tab.n[tab.n$Sequence==1 & tab.n$Period==1,]$x
n2<-tab.n[tab.n$Sequence==2 & tab.n$Period==1,]$x
n<-n1+n2



####-----carryover effect -----------##
uik<-aggregate(chow$lnAUCT, list(Sequence=chow$Sequence, Subject=chow$$
colnames(uik)<-c("Sequence","Subject","uik")
muk<-aggregate(uik$uik, list(Sequence=uik$Sequence),mean)
colnames(muk)<-c("Sequence","muk")
print(muk)
hatc<-muk[2,2]-muk[1,2]
hatc
du<-merge(uik,muk)
sigu2<-sum((du$uik-du$muk)^2/(n1+n2-2))
se.sigu<-sqrt(sigu2*((1/n1)+(1/n2)))
TC<-hatc/se.sigu
```

```
TC
pc<-2*(1-pt(abs(TC),n1+n2-2))
pc
###----formulation effect-----###
dik<-aggregate(chow$lnAUCT,list(Subject=chow$Subject,Sequence=chow$Se
dik$x<-dik$x/2
colnames(dik)<-c("subject","Sequence","dik")
mdk<-aggregate(dik$dik,list(Sequence=dik$Sequence),mean)
colnames(mdk)<-c("Sequence","mdk")
hatF<-mdk[1,2]-mdk[2,2]

dF=merge(dik,mdk)
sigd2<-sum((dF$dik-dF$mdk)^2)/(n1+n2-2)
sigd2
se.sigd<-sqrt(sigd2*((1/n1)+(1/n2)))
TF<-hatF/se.sigd
TF
pf<-2*(1-pt(abs(TF),n1+n2-2))
pf


mdrug<-tapply(chow$lnAUCT,list(Formulation=chow$Formulation),mean)
ybarT<-mdrug["T"]
ybarR<-mdrug["R"]


##-----Confidence interval-----##
alphaCI<-0.1
qt.alpha<-qt(1-alphaCI,n1+n2-2)
#sigd2<-anova(mdlnAUCT)[5,3]/2
low1<-(ybarT-ybarR)-qt.alpha*sqrt(sigd2)*sqrt(((1/n1)+(1/n2)))
```

```
up1<-(ybarT-ybarR)+qt.alpha*sqrt(sigd2)*sqrt(((1/n1)+(1/n2)))
cat("The classical CI1=(",round(low1,6),",",round(up1,6),")",sep="","
low2<-((low1/ybarR)+1)*100
up2<-((up1/ybarR)+1)*100
cat("The Ratio CI2=(",round(low2,5),",",round(up2,5),")",sep="","\n\n



##---westlake---##
mdrug<-tapply(chow$lnAUCT,list(Formulation=chow$Formulation),mean)
ybarT<-mdrug["T"]
ybarR<-mdrug["R"]


k12<-2*(ybarR-ybarT)/sqrt(sigd2*(((1/n1)+(1/n2))))
k2<-uniroot(function(k2)pt(k12-k2,n1+n2-2)-pt(k2,n1+n2-2)-(1-alphaCI)
k1<-k12-k2
cat("the westlake k1=",k1," and k2=",k2,sep="","\n\n")
low.west<-k2*sqrt(sigd2*(1/n1+1/n2))-(ybarR-ybarT)
up.west<-k1*sqrt(sigd2*(1/n1+1/n2))-(ybarR-ybarT)
cat("The westlake CI for mu_T-Mu_A is (",low.west,",",up.west,",",sep


##--two 0ne-sided---##
theta.L<- -0.2*ybarR
theta.U<- 0.25*ybarR
#theta.U<- 0.25*ybarR
TL<-(ybarT-ybarR-theta.L)/sqrt(sigd2*((1/n1)+(1/n2)))
TU<-(ybarT-ybarR-theta.U)/sqrt(sigd2*((1/n1)+(1/n2)))
```