Characterising tuberculosis treatment success and failure using metabolomics

F Kamfer

21090424

Dissertation submitted for the degree *Magister Scientiae* in Biochemistry at the Potchefstroom Campus of the North-West University

Supervisor: Prof DT Loots

September 2013

Characterising tuberculosis treatment success and failure using metabolomics

Fanie Kamfer

Honns. B.Sc. (Biochemistry)

Project submitted in partial fulfilment of the requirements for the degree

Master of Science

in

Biochemistry

at the

School for Physical and Chemical Sciences, Centre for Human Metabonomics,

NORTH-WEST UNIVERSITY (POTCHEFSTROOM CAMPUS)

Supervisor: Prof Du Toit Loots

Potchefstroom

2013

ACKNOWLEDGEMENTS

The contributions of the following institutions and individuals made towards the successful completion of this study are hereby acknowledged:

Prof. Du Toit Loots, my study supervisor, thank you for your guidance and motivation during the course of this M.Sc. study.

The DST/NRF Centre of Excellence for Biomedical TB research, Stellenbosch University for providing the patient collected urine samples used in this investigation.

The National Research Foundation (NRF) of South Africa, the Technology Innovation Agency (TIA) and the North-West University (NWU) for the research grants provided.

I would also like to thank my friends and all the students and staff members of the Centre for Human Metabonomics, for all the laughter, support and advice you've provided me with over the years.

My parents, Elrie and Santa, for their ongoing physical, emotional and financial support during the course of this M.Sc. study.

Finally, I want to thank my heavenly Father for guiding me to this point in my live. Thank You for always protecting me, giving me strength, and for all the blessings you bestowed upon me.

"I can do everything though Christ who gives me strength."
Philippians 4:13

SUMMARY

Tuberculosis (TB) is one of the deadliest infectious diseases of our time, with 1.4 million deaths globally, recorded in 2010 (3800 deaths a day) by the World Health Organization (WHO). Currently, South Africa ranks third on the 2011 list of 22 high-burden TB countries in the world and it was estimated that each active-TB person could potentially infect 10–15 people annually. The WHO additionally reported that in the year 2009, 87% of all TB patients worldwide were successfully treated, with a treatment success rate of 74% reported for South Africa. Despite this however, non-adherence to anti-TB treatment is still a major issue, due to it resulting in a global increased prevalence of drug resistant TB and subsequently TB treatment failure. Treatment failure is thought to be caused by a number of factors, however, it still remains largely misunderstood. One aspect of this, that isn't clearly addressed in the literature, is the underlying variation in each patient, resulting in his/her varying reaction to the drug regimen, and hence it's varying efficacy from one patient to the next. Furthermore, little is known about the underlying variation of the host to the primary TB infection or response to the TB disease state, and how some patients have more effective mechanisms for eliminating the infection, or recovering from the disease.

Considering this, a metabolomics research study using GC×GC-TOFMS was conducted, in order to identify potential metabolite markers which may be used to better characterise the underlining mechanisms associated with poor treatment outcomes (treatment failure).

The first aim was to evaluate the accuracy and efficiency of the methodology used, as well as to determine the capability and accuracy of the analyst to perform these methods. In order to evaluate the GCxGC-TOFMS analytical repeatability, one QC sample was extracted and injected repeatedly (6 times) onto the GC×GC-TOFMS. Similarly, the analyst's repeatability for performing the organic acid extraction and analyses was also determined, using 10 identical QC samples, which were extracted and injected separately. CV values were subsequently calculated from the collected and processed data as a measure of this. Of all the compounds detected from the 6 QC sample repeats used for GCxGC-TOFMS repeatability, 95.59% fell below a 50% CV value, and 93,7% of all the compounds analysed for analyst repeatability had a CV < 50.

Subsequently, using the above metabolomics approach, in addition to a wide variety of univariate and multivariate statistical methods, two patient outcome groups were compared. A sample group cured from TB after 6 months of treatment was compared vs a sample group where treatment failed after the 6 month period. Using urine collected from these two patient groups at various time points, the following metabolomics comparisons where made: 1) at time of diagnosis, before any anti-TB treatment was administrated, 2) during the course of treatment, in order to determine any variance in these groups due to a varying response to the anti-TB drugs, 3) over the duration of the entire 6 months treatment regimen, in order to determine if differences exist between the two groups over time.

A clear natural differentiation between the cured and failed outcome groups were obtained at time of diagnosis, and a total of 39 metabolites markers were subsequently identified. These metabolites were classified according to their various origins, and included (1) those associated with the presence of *M. tuberculosis* bacteria, (2) those resulting from an altered host metabolism due to the TB infection, and (3) metabolites of various exogenous origins. The detailed interpretation of these metabolites suggests that a possible underlying RCD or some sort of mitochondrial dysfunction may be present in the treatment failure group, which may also be induced through an external stimulus, such as alcohol consumption. We hypothesise that this may possibly result in a far greater severity to *M. tuberculosis* infection in this group, subsequently causing a reduced capacity for a successful treatment outcome, also considering the critical role of the mitochondria in the metabolism of anti-TB drugs.

Furthermore, 20 metabolite markers were identified when comparing the two outcome groups during the treatment phase of this metabolomics investigation. A vast majority of these 20 metabolites were also identified as markers for time 0 (time of diagnosis). Additionally, metabolites associated with anti-TB drug induced side effects, were also found to be comparatively increased in the treatment failure group, indicative of more pronounced liver damage, accompanied by metabolites characteristic of a MADD metabolite profile, due to a deficient electron transport flavoprotein, confirming previous experiments done in rats. These side effects have also previously been implicated as a major contributor of poor treatment compliance, and ultimately treatment failure.

Lastly, 35 metabolite markers were identified by time dependent statistical analysis and represented those metabolites best describing the variation between the treatment outcome

groups over the entire study duration (from diagnosis, to week 26). This time dependent statistical analysis identified markers, using an alternative statistical approach, and confirmed previous findings and added in a better characterisation of treatment failure.

Considering the above, we successfully applied a metabolomics approach for identifying metabolites which could ultimately aid in the prediction and monitoring of treatment outcomes. This additionally led to a better understanding and or characterisation of the phenomenon known as treatment failure, as well as the underlying mechanisms related to this occurrence.

Keywords: tuberculosis; metabolomics; treatment failure; metabolite marker; GCxGC-TOFMS

Opsomming

Tuberkulose (TB) is een van die dodelikste aansteeklike siektes van ons tyd. In 2010 het die Wêreld Gesondheidsorganisasie (WHO) 1,4 miljoen sterftes wêreldwyd aangeteken (3800 sterftes per dag). Suid-Afrika beklee in 2011 die derde posisie op die lys van 22 lande wat die hoogste aantal TB gevalle ter wêreld het. Daar word beraam dat elke TB-aktiewe persoon, potensiëel die siekte kan oordra aan 10-15 mense per jaar (WHO, 2011 (b)). Die WHO het ook berig dat in 2009, 87% van alle TB-pasiënte wêreldwyd suksesvol behandel is, met 'n behandelingsukseskoers van 74% vir Suid-Afrika (WHO, 2011 (b)). Ten spyte van bogenoemde, wek die onvermoë om anti-TB behandeling te voltooi nog steeds groot kommer en het aanleiding gegee tot 'n wêreldwye toename in die voorkoms van aangemelde veelvoudige-middel-weerstandbiedende TB (MDR-TB), wat uiteindelik uitloop op onsuksesvolle behandeling.

Onsuksesvolle behandeling kan toegeskryf word aan verskeie faktore en is nog steeds ‘n onderwerp wat baie aandag moet geniet. Een aspek wat nie deeglik bespreek word in die literatuur nie, is die onderliggende variasie in elke pasiënt, wat daartoe lei dat sommige pasiënte anders reageer op anti-TB medikasie as ander. Verder is min inligting bekend rakende die onderliggende variasie van die gasheer tot die primêre TB-infeksie of 'n reaksie op die TB siektetoestand en hoe sommige pasiënte meer effektiewe meganismes in plek het vir die effektiewe eliminering van die TB infeksie. Deur bogenoemde in ag te neem, is daar van ‘n GCxGC-TOFMS gebaseerde metabolomika navorsingsbenadering gebruik gemaak om potensiële metabolietmerkers te identifiseer wat gebruik kan word om die onderliggende meganismes wat verband hou met die voorkoms van mislukte TB behandeling onder pasiënte beter te karakteriseer.

Die eerste doelwit van die navorsingsbenadering was om die akkuraatheid en doeltreffendheid van die metodologie wat gebruik is, te evalueer, asook om die vermoë en die akkuraatheid van die analis vas te stel met betrekking tot die analitiese metodes wat in die studie uitgevoer is. Ten einde die GCxGC-TOFMS se analitiese herhaalbaarheid te evalueer, is een kwaliteitskontrole (QC) urienmonster geëkstraheer en herhaaldelik (6 keer) in die GCxGC-TOFMS ingevoer. Net so is die analis se herhaalbaarheid vir die uitvoer van die organiese suurekstraksie bepaal, deur gebruik te maak van 10 QC monsters wat

geëkstraheer en afsonderlik in die GCxGC-TOFMS ingevoer is. Vervolgens is CV waardes van die versamelde en verwerkte data bereken, om as 'n maatstaf te dien vir die bepaling van die herhaalbaarheid. Van al die verbindings wat geëkstraheer is van die 6 QC monster wat gebruik is vir GCxGC-TOFMS herhaalbaarheid, het 95,59% ‘n CV waarde kleiner as 50% getoon, en 93,7% van al die verbindings geëkstraheer uit die QC monsters vir die bepaling van die analis herhaalbaarheid, het ‘n CV waarde kleiner as 50 gehad.

Daarna, deur gebruik te maak van bogenoemde metabolomika benadering, is 'n wye verskeidenheid van enkelveranderlike en meerveranderlike statistiese metodes gebruik om uiteindelik te kan onderskei tussen pasiënte wat genees is na die 6 maande tydperk van TB behandeling, teenoor diegene wat nie genees is nie. Die twee pasiëntgroepe is op verskeie punte gedurende die 6 maande behandelingsiklus met mekaar vergelyk: 1) tyd van diagnose, voordat enige anti-TB medikasie ingeneem is, 2) gedurende die duur van die behandeling, ten einde enige variasie in hierdie groepe te bepaal as gevolg van 'n wisselende reaksie op die anti-TB-medikasie, 3) groepe is oor die hele 6 maande behandelingstydperk vergelyk, ten einde te bepaal of daar met verloop van tyd enige verskille bestaan tussen die twee groepe.

'n Duidelike natuurlike differensiëring tussen die suksesvolle en onsuksesvolle behandelingsuitkomsgroepe is verkry op die diagnosetydstip voor enige anti-TB medikasie geneem is. 'n Totaal van 39 metaboliete merkers is vervolgens geïdentifiseer. Hierdie metaboliete is geklassifiseer in drie hoofgroepe, insluitend (1) die wat geassosieer word met die teenwoordigheid van *M. tuberculosis* bakterieë, (2) die wat voorkom as gevolg van 'n veranderde gasheermetabolisme vanweë die TB-infeksie, en (3) metaboliete van verskeie eksogene oorsprong. Die gedetailleerde interpretasie van hierdie metaboliete dui daarop dat 'n moontlike onderliggende respiratoriese kettingdefek of mitokondriale wanfunksionering teenwoordig mag wees in die onsuksesvolle behandelingsgroep. Daar word gespekuleer dat dit aanleiding kan gee tot ‘n erger infeksie, asook die gevolglike verlaagde kapasiteit vir ‘n suksesvolle behandelingsuitkoms, gedeeltelik weens die feit dat die mitokondria ‘n belangrike rol speel in anti-TB medikasiemetabolisme.

Gedurende die hele behandelingsfase is 20 metabolietmerkers geïdentifiseer tussen die suksesvolle en onsuksesvolle behandelingsgroepe van hierdie metabolomika studie. Die oorgrote meerderheid van hierdie 20 metaboliete het ooreenstemming getoon met die merkers, wat tydens die diagnosetydstip geïdentifiseer is. Addisioneel is metaboliete wat

verband hou met newe-effekte wat deur anti-TB medikasie geïnduseer word, in hoë konsentrasie geïdentifiseer in die onsuksesvolle behandelingsgroep. Metaboliete wat ‘n aanduiding gee van lewerskade sowel as metaboliete wat kenmerkend is aan 'n Meervoudige Asiel-KoA Dehidroginase defek (MADD) metaboliete profiele, is ook waargeneem in die onsuksesvolle behandelingsgroep. Hierdie newe-effekte is ook voorheen beskou as 'n groot bydraer tot onsuksesvolle behandelingsuitkomste, vanweë TB pasiënte wat die neem van medikasie staak. Ten slotte is 35 metabolietmerkers geïdentifiseer deur tydafhanklike statistiese analise en verteenwoordig die metaboliete wat die beste beskrywing gee van die variasie tussen die twee behandelingsuitkomsgroepe oor die hele tydperk (van diagnose, tot week 26). Die statistiese metode is gebruik ten einde vorige bevindings te bevestig sowel as om addisionele merkers te identifiseer wat verder kan bydra om suksesvolle en onsuksesvolle behandelingsuitkomste beter te karakteriseer.

Deur bogenoemde in ag te neem, is dit duidelik dat ons daarin geslaag het om die metabolomikabenadering suksesvol toe te pas deur metaboliete te identifiseer wat sal help om die behandelingsuitkoms van TB pasiënte vroeër te bepaal en beter te monitor. Verder sal dit ook lei tot 'n beter begrip en karakterisering van die verskynsel wat bekend staan as onsuksesvolle behandeling, sowel as die onderliggende meganismes wat verband hou met hierdie verskynsel.

Sleutel woorde: tuberkulose; metabolomika; onsuksesvolle behandeling; metaboliet Merker; GCxGC-TOFMS.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

SUMMARY ii

KEY WORDS iv

OPSOMMING v

SLEUTEL WOORDE vii

LIST OF ABREVIATIONS xiii

LIST OF FIGURES xvi

LIST OF TABLES xvii

CHAPTER 1: Introduction 1

1.1 INTRODUCTION 2

1.2 STRUCTURE OF DISSERTATION 3

1.3 AUTHOR CONTRIBUTIONS 5

CHAPTER 2: Literature study 7

2.1 BACTERIOLOGY 8

2.2 PATHOPHYSIOLOGY 8

2.3 TUBERCULOSIS TREATMENT 10

2.3.1 First line anti-TB drugs 12

2.3.1.1 Isoniazid 12

2.3.1.2 Rifampicin 14

2.3.1.3 Pyrazinamide 15

2.3.1.4 Ethambutol 17

2.4 TUBERCULOSIS TREATMENT FAILURE 19

2.4.1 Non-adherence to treatment and resistance to anti-TB drugs 19

2.4.2 Poverty, poor lif quality and limeted availability of anti-TB drugs 21

2.4.3 Lack of patient knowlege and physician mismanagement 21
2.4.4 Other disease states and contributing factors 22
2.5 TUBERCULOSIS DIAGNOSTICS 23
2.5.1 Technigues for the diagnosis of *M. tuberculosis* 24
2.5.1.1 Tuberculin skin test 24
2.5.1.2 Interferon Gamma Release Assays 24
2.5.2 Active TB diagnostic techniques 25
2.5.2.1 Chest radiology 25
2.5.2.2 Microscopy techniques 26
2.5.2.3 Bacteriological cultures 26
2.5.2.4 Nucleic acid amplification techniques 27
2.5.2.5 Serology 28
2.5.2.6 Phage assay 28
2.5.2.6 Xpert MTB/RIF assay 29
2.6 Predicting treatment outcome 30
2.6.1 Predicting treatment outcome by sputum smear results 30
2.6.2 Predicting treatment outcome by anti-TB drug resistance 30
2.6.3 Predicting treatment outcome by monitoring bodyweight variation 31
2.7 Patient collected sample material used for characterising and daignosing TB 32
2.7.1 Blood 32
2.7.2 Sputum 33
2.7.3 Urine 34
2.8 Metabolomics 36
2.8.1 TB Metabolomics research to date 38

CHAPTER 3: Aims and objectives 40

3.1 PROBLEM STATEMENT 41
3.2 AIMS 42

3.3 OBJECTIVES 42

CHAPTER 4: Materials and methods 44

4.1 REAGENTS 45

4.2 URINE SAMPLES 45

4.3 QUALITY CONTROL SAMPLES 47

4.4 METHODS 48

4.4.1 Creatinine determination 48

4.4.2 Extraction and derivatization of organic acids 49

4.4.3 Gas chromatography mass spectrometry analysis 50

4.5 DATA PROCESSING 51

4.5.1 Deconvolution and peak identification 51

4.5.2 Peak alignment 52

4.5.3 Data pre-processing 52

4.6 STATISTICAL DATA ANALYSIS 55

4.6.1 Univariate analysis 55

2.6.1.1 *t*-test 56

2.6.1.2 Effect size 56

2.6.1.3 Analysis of variance 57

4.6.2 Multivariate analysis 58

2.6.2.1 Principal component analysis 58

2.6.2.2 Partial-least squares discriminant analysis 59

2.6.2.3 ANOVA-simultaneous component analysis 59

4.7 SUMMARY OF EXPERIMENTAL DESIGN 60

CHAPTER 5: Results and discussion 63

5.1 REPEATABILITY AND RELIABILITY OF DATA 64

5.1.1 GCxGC-TOFMS and analyst repeatability 64

5.1.2 Batch effect 68

5.2 MATERIALS AND METHODS 70

5.2.1 Metabolomics comparisons of successful and unsuccessful treatment outcomes at time of diagnosis 70

5.2.1.1 PCA differentiation between successful and unsuccessful treatment outcomes at time of diagnosis 71

5.2.1.2 Metabolite marker identification 72

5.2.1.2.1 Markers associated with *M. tuberculosis* metabolism 78

5.2.1.2.2 Markers associated with an altered host metabolism 78

5.2.1.2.3 Markers of exogenous origens 83

5.2.2 Metabolomics comparisons of successful and unsuccessful treatment outcomes during treatment 84

5.2.2.1 PCA differentiation between successful and unsuccessful treatment outcomes during treatment 85

5.2.2.2 Metabolite marker identification 86

5.2.2.2.1 Markers associated with an altered human metabolism 89

5.2.2.2.2 Markers associated with anti-TB drug induced side effects 92

5.2.2.2.3 Markers of exogenous origens 92

5.2.3 Metabolomics comparisons of successful and unsuccessful treatment outcomes over time 94

5.2.3.1 ANOVA and ASCA statistical analysis between successful and unsuccessful treatment outcome groups 95

5.2.3.2 Metabolite marker identification 96

5.2.3.2.1 Markers associated with host response to TB and futher aggravated by anti-TB drugs 98

5.2.3.2.2 Metabolites associated with anti-TB drug induced side effects 102

5.2.3.2.3 Markers of exogenous origens 103

Chapter 6: Conclusion 106

6.1 General conclusion.....107

6.2 Future recommendation 111

Chapter 7: References 113

APPENDIX 1: 123

APPENDIX 2: 126

LIST OF ABREVIATIONS

2D	- Two dimensional
°C	- Degrees Celsius
AGV	-Average
AIDS	- Acquired Immune Deficiency Syndrome
ANOVA	-Analysis of variation
Anti-TB	- Anti-tuberculosis
ASCA	-ANOVA-simultaneous component analysis
Aspergillus sp.	- *Aspergillus species*
ATP	- Adenosine Triphosphate
BCG	- Bacille Calmette-Guerin
BMI	- Body Mass Index
bp	- Base pair(s)
BSTFA	- Bis(trimethylsily)-trifluoracetamide
CDC	- Centre for Disease Control and Prevention
CE	- Capillary Electrophoresis
CFP-10	- Culture Filtrate Protein
CoA	- Coenzyme A
Cum	- Cumulative
CV	- Coefficients of Variation
CYP450	- Cytochrome P450
DNA	- Deoxyribonucleic Acid
DOTS	- Directly Observed Treatment, Short-Courses
ELISA	- Enzyme-Linked Immunosorbent Assay
ELISpot	- Enzyme-Linked Immunospot Assay
EMB	- Ethambutol
ES	- Effect Sizes
ESAT-6	- Early Secreted Antigenic Target
ETC	- Electron Transport Chain
ETF	- Electron-Transfer Flavoprotein
ETF-QC	- Electron-Transfer Flavoprotein: Ubiquinone Oxidoreductase
eV	- Electron Volt
FDA	- Food and Drug Administration
g	- Gram
GABA	- γ-ammino-bytyric acid
GC	- Gas Chromatography
GCxGC-TOFMS	- Two dimensional Gas Chromatography, coupled with a Time of Flight Mass Spectrometer
GGC	-γ-Glutamyl cycle
H_2O_2	- Hydrogen Peroxide
HCl	- Hyrdochloric Acid
HIV	- Human Immunodeficiency Virus
HOCl	- Hypochlorous Acid

HPHPA	- 3-(3-hydroxyphenyl)-3-hydroxypropapoinc Acid
IDO-1	- Indolamine-2,3-dioxygenase-1
IFN-γ	- Interferon Gamma
IGRA	- Interferon Gamma Release Assay
INH	- Isoniazid
InhA gene	- Encodes an enoyl-acyl carrier protein reductase
katG gene	- Encodes the catalase-peroxidase enzyme in *M. tuberculosis*
LAM	- Lipoarabinomannan
LC	- Liquid Chromatography
m	- Meter
M. bovis	- *Mycobacterium bovis*
M. tuberculosis	- *Mycobacterium tuberculosis*
m/z	- Mass-to-Charge Ratio
mAGP	- Mycolyl arabinogalactan-peptidogalactan
MANOVA	-Multivariate ANOVA
Max	-Maximum
MDR-TB	- Multi-Drug-Resistant tuberculosis
Min	-Minimum
mL	- Millimeter
MPS	- Multi-Purpose Sampler
MS	- Mass Spectrometry
N	- Mol per millilitre (mol/mL)
NA	-Not Applicable
NAA	-Nucleic Acid Amplification
Na_2SO_4	- Sodium Sulfate
NAA	- Nucleic Acid Amplification
NAD+	- Nicotinamide Adenine Dinucleotide
NADPH	- Nicotinamide Adenine Dinucleotide Phosphate-Oxidase
NAT 2	- N-acetyl Transferase 2
NMR	- Nuclear Magnetic Resonance
NWU	- North-West University
O_2^-	- Superoxide
$ONOO^-$	- Peroxynitrite
PC	- Principal Component
PCA	- Principal Component Analysis
PCR	- Polymerase Chain Reaction
PDC	-Pyruvate dehydrogenase
PKU	- Phenylketonuria
PLS-DA	- Partial Least Squares – Discriminant Analysis
pncA gene	- Encodes pyrazinamidase
POA	- Pyrazinoic Acid
PPD	- Purified Protein Derivative
PZA	- Pyrazinamide
PZase	- Pyrazinamidase

QC	- Quality Control
QC-CV	- Quality Control – Coefficient of Variation
RCD	- Respiratory chain deficiency
RD	- Refsum's disease
RIF	- Rifamycins
RNA	- Ribonucleic Acid
RNS	- Reactive Nitrogen Species
ROS	- Reactive Oxygen Species
rpm	- Rounds per Minute
rpoB gene	- Encodes DNA-dependent RNA-polymerase of M. tuberculosis
RpsA	- Ribosomal Protein S1
s	- Second(s)
S/N	- Signal to Noise Ratio
SPE	-Squared prediction error
STDEV	-Standard deviation
TB	- Tuberculosis
TCA cycle	- Tricarboxylic Acid Cycle
TMCS	- Trimethylchlorosilane
TMIC	- The Metabolomics Innovation Centre
TOFMS	- Time Of Flight Mass Spectrometry
TST	- Tuberculin Skin Test
VIP	- Variable Influence on the Projection
WHO	- World Health Organization
XDR-TB	- Extensively Drug-Resistant Tuberculosis
µL	- Microliter
µm	- Micrometer

LIST OF FIGURES

Figure 4.1: Summary of the urine sample collection time intervals 47

Figure 4.2: Schematic representation of the experimental design used for this metabolomics study 61

Figure 4.3: Work-flow of the statistical methods used in this metabolomics study.....62

Figure 5.1: Distribution of the coefficients of variation (CV) values determined from the relative concentrations of all the compounds detected using the QC samples as a measure for a) GCxGC/TOFMS repeatability and b) Extraction method/ Analyst repeatability, subsequent to an organic acid extraction, and GCxGC/TOFMS analysis.....67

Figure 5.2: 2D score plots of a) PC 1 vs. PC 2, b) PC 2 vs. PC 3 and c) PC 1 vs. PC 3 of the 10 batches in which the 210 patient urine samples were analyzed. The percentage cumulative variance explained by the first two PCs (R2X cum) was 13.4%, of which PC 1 explained 7.0% and PC 2 explained 5.4%.....69

Figure 5.3: PCA scores plot illustrating PC1 vs. PC2 of the successful and failed treatment outcome group at time of diagnosis, subsequently to organic acid extraction and GCxGC/TOFMS analysis, indicating a differentiation of the two outcome groups. The variance explained by each PC is given in parenthesis.....72

Figure 5.4: Statistical approach used to identify metabolite markers which best explain the variation between the successful and failed outcome at time of diagnosis.....74

Figure 5.5: Schematic representation of elevated concentrations of the various metabolites found in the urine of the treatment failure group. This diagram indicates the metabolites associated with the M. tuberculosis infection, an altered host metabolism due to the infection and metabolites of exogenous origin..... 77

Figure 5.6: PCA scores plot illustrating PC1 vs. PC2 using data calculated by determining the difference in metabolite concentrations at week 4 of treatment and week zero (before treatment), subsequently to organic acid extraction and GCxGC/TOFMS analysis, indicating a differentiation of the two outcome groups. The variance explained by each PC is given in parenthesis 86

Figure 5.7: Statistical approach used to identify metabolite markers which best explain the variation between the successful and failed outcome during treatment.....87

Figure 5.8: Major patterns associated with the interaction between time (0, 1, 2, 4, 6 representing week 0, 1, 2, 4 and 26) and treatment outcome (Indicated as cured and failed) as identified by ASCA analysis (figures generated on MetaboAnalyst).....96

LIST OF TABLES

Table 1.1: Research team…….....6

Table 2.1: Official WHO definitions of TB treatment outcomes.....23

Table 2.2: Potential advantages and disadvantages of diagnostic material.....35

Table 4.1: Volume of urine used in the organic acid extraction method, according to creatinine values.....49

Table 4.2: Formulas used to determine the needed volume of reagents used in the organic acid extraction method.....49

Table 4.3: Summary of the total aligned compounds for each time interval, as well as the amount of compounds which could be annoted, or otherwise remained unidentified.....55

Table 5.1: Coefficients of variation (CV) values, calculated for ten compounds with the use of compound area and the relative concentration respectively, after GCxGC-TOFMS analysis. …….....65

Table 5.2: Coefficients of variation (CV) values, calculated for ten compounds detected in all QC samples after organic acid extraction analysis and derivatization, with the use of compound area and concentration respectively.....66

Table 5.3: Metabolite markers best describing the variation between the successful and faild treatment outcome groups.....75

Table 5.4: Metabolite markers best describing the variation between the successful and faild treatment oucome groups during the standard TB treatment regimen.....88

Table 5.5: Metabolite markers best describing the variation between the successful and faild treatment oucome groups from time of diagnosis till the end of the six month treatment regimen.....97

Table 1: Clinical and microbial information of the anonyms patients of which urine samples were collected and used in this metabolomics study…….....124

Chapter 1

Introduction

1.1 Introduction

Tuberculosis (TB) is one of the deadliest infectious diseases of our time, with 1.4 million deaths globally, recorded in 2010 (3800 deaths a day) by the World Health Organization (WHO, 2011). TB is also considered the leading cause of death in patients infected with the human immunodeficiency virus (HIV). This is not surprising, considering the infection rate of 8.8 million new TB cases reported for the same year (WHO, 2011(b)). It has also been reported that Africa accounts for the majority of TB related deaths (71%) in patients co-infected with HIV, with South Africa alone, contributing to an astonishing 31% of these deaths (WHO, 2011). Although the TB incidence has been shown to decline in many third world countries, this occurrence is quite the opposite for the many poor developing countries in the world (González-Martín *et al.,* 2010). Currently, South Africa ranks third on the 2011 list of 22 high-burden TB countries in the world and it was estimated that each active-TB person could potentially infect 10–15 people annually (WHO, 2011(b)).

The increase of TB prevalence in these countries, may be attributed to the increased HIV incidence, anti-tuberculosis (anti-TB) drug resistance, travel and immigration from other higher TB prevalence countries and increased poverty associated with overcrowding and malnutrition (Knechel, 2009). On the upside, however, the WHO had reported recently a decrease of 80% in the global TB death rate from 1990 to 2010. The WHO additionally reported that in the year 2009, 87% of all TB patients worldwide were successfully treated, with a treatment success rate of 74% reported for South Africa (WHO, 2011(b)). Despite this however, non-adherence to treatment is still a major issue, due to it resulting in a global increased prevalence of drug resistant TB and subsequently TB treatment failure. In 2010, an estimated 650 000 cases of multi-drug resistant TB (MDR-TB) were reported globally, which lead to an increase of 46 000 TB patients in need of MDR-TB treatment (Parida & Kaufmann, 2009; WHO, 2011(a)). In South-Africa, MDR-TB cases have increased at an astonishing rate from 2000 cases in 2005 to 7350 cases in 2007. An increase of extensively drug-resistant TB (XDR-TB) has also been observed which has increased from 74 cases in 2004 to 536 cases in 2007 (WHO, 2011(b)). Clearly, greater efforts are required to strengthen the overall TB control in order to lower these TB incidences.

Treatment failure is thought to be caused by a number of factors, however, it still remains largely misunderstood. One aspect of this, that isn't clearly addressed in the literature, is the underlying variation in each patient, resulting in his/her varying reaction to the drug regimen, and hence it's varying efficacy from one patient to the next. Furthermore, little is known about the underlying variation of the host to the primary TB infection or response to the TB disease state, and how some patients have more effective mechanisms for eliminating the infection, or recovering from the disease. Many factors may contribute to this, including diet, social practices (drinking and smoking), and perhaps even genetics. These mechanisms are poorly understood, and very little research has been done on treatment failure from this perspective. This metabolomics study however, may shed light on treatment failure from this perspective, as it focuses on the response of the human metabolome to the TB infection or disease, and the medication used for treating this, comparing those individuals with a positive treatment outcome (cured) to those in whom the treatment approach failed (treatment failure).

Considering this, we conducted a metabolomics research study using GC×GC/TOFMS in order to identify potential metabolite markers which may be used to better characterise the underlining mechanisms associated with poor treatment outcomes (treatment failure). This in turn may provide valuable information regarding treatment outcome which will aid in development of alternative diagnostic and treatment approaches, which will serve well in lowering the incidence rate of TB.

1.2 Structure of dissertation

This dissertation is a compilation of chapters written specifically to comply with the requirements of the North-West University, Potchefstroom campus, for the completion of the degree Master of Science (Biochemistry) in dissertation format.

The current chapter, Chapter 1, gives a brief background to the conducted study, focusing on the epidemiology reports pertaining to TB, especially which relate to poor treatment outcomes (treatment failure), justifying the need for this metabolomics investigation.

This chapter also discusses the structure of the dissertation and clarifies the contributions of each individual of the research team, towards the execution and completion of this study.

In Chapter 2, a summary of TB in general is provided, as well as related aspects, required as a basis for a better understanding of this epidemic and the results and discussions that follow. Due to the fact that metabolomics as a research discipline is relatively new, an explanation of what metabolomics is and its applications for characterising TB are also addressed in this section.

Chapter 3 provides the problem statement, aims and objectives of this metabolomics study.

The experimental design, sample collection procedures and the metabolomics research methodology, including data processing and statistical analysis methods used, are described in Chapter 4.

Chapter 5 describes the results obtained using the metabolomics research methodology and statistical analysis described in Chapter 4, in order to differentiate between the successful and unsuccessful treatment outcome patient sample groups collected at the various time intervals. Before application of the research methodology to answering the biological question however, the method was validated, by determining the repeatability and reliability of the data collected using the approach described in Chapter 4. Subsequently, this approach was used to determine if differentiation, on the basis of the GCxGC-TOFMS analysed urinary metabolic profiles of these patient groups, could be achieved at (1) time of diagnosis, (2) at a time point during treatment protocol and (3) globally, considering the whole treatment period, as a means for identifying markers better explaining the phenomenon of treatment failure. Finally this Chapter includes a comprehensive discussion on the identified metabolite markers best describing the variation between the outcome groups at the above mentioned time points, focussing on their various origins, and how they can be used for better understanding the phenomenon of treatment failure.

Chapter 6 is the final conclusion, considering all the results obtained, in addition to recommendations for future research, potentially emanating from this study.

1.3 Author contribution

The principle author of this thesis is Fanie Kamfer. The contribution of the co-authors, co-workers and collaborators made towards this work is given in Table 1.1.

The following is a statement from the co-authors confirming their individual roles in the study and giving their permission that the data generated and conclusions made may form part of this thesis.

I declare that my role in this study, as indicated in Table 1.1, is representative of my actual contribution and I hereby give my consent that this work may be published as part of the M.Sc. dissertation of Fanie Kamfer.

.....

Prof. D.T. Loots

.....

F. Kamfer

Table 1.1: Research team

Co-author	Co-worker	Collaborator	Contribution
F. Kamfer (B.Sc. Honns. Biochemistry)			Responsible, together with the promoter, for the planning, execution, data analyses, and writing of the thesis, publication, and all other documentation and experimentation associated with this study.
Prof. D.T. Loots (Ph.D. Biochemistry)			Promoter: Co-ordinated and supervised all aspects of the study including: study design, planning, execution, and the writing of the thesis, publication, and all other documentation and experimentation associated with this study.
	M van Reenen (M.Sc. Statistics)		Assisted with study design and data analyses from a statistical perspective.
		NRF/DST Centre of Excellence in Biomedical Tuberculosis Research, Faculty of Health Sciences, University of Stellenbosch	Provided the patient urine samples used in this metabolomics study
		AMPATH laboratories	Determined Creatinine values of all patient urine samples collected

Chapter 2

Literature Study

2.1 Bacteriology

The pathogen *M. tuberculosis*, the causative agent of tuberculosis, is a small, non-motile, aerobic bacillus usually located in the upper respiratory tract.

The cell wall of *M. tuberculosis* is considered to have a unique structure, consisting of a substantial amount of mycolic acids. Mycolic acids are long chain methyl branched fatty acids characteristic to *Mycobacterium,* and are covalently linked to an underlining peptidoglycan bound polysaccharide, arabinogalactan, to form a unique lipid barrier around the organism (Knechel, 2009). This barrier increases the resistance and inpermeability of the organism against antimicrobial agents, acidic and alkaline compounds, osmotic lysis, lysozymes and immune defence mechanisms (Willey, 2008). The bacteria's growth rate and virulence are also influenced by the content of its cell wall, and lipoarabinomannan (a structural carbohydrate antigen) is thought to be responsible for the survival of this organism against host macrophages (Knechel, 2009).

M. tuberculosis is classified as a gram positive, acid-fast bacteria. This characteristic is also attributed to the high lipid content of its cell wall. The mycolic acids prevent conventional dyes from binding to their cell walls, however, harsher staining methods, which drive the dyes into these cells using heat and phenol (the Ziehl-Neelsen method) have been shown effective for staining these organisms. Consequently, this method is most widely used for identifying these organisms, and for disease diagnostics (Willey *et al.,* 2008).

2.2 Pathophysiology

Transmission of TB occurs when an individual with active pulmonary or laryngeal TB sneezes, coughs or speaks, releasing infectious air droplet nuclei into the environment (Knechel, 2009). These *M. tuberculosis* containing nuclei are typically 1-5 μm in diameter, and may remain airborne for long periods of time (Frieden *et al*., 2003).

Infection may occur once these droplets are inhaled and enter the respiratory tract of the infected host. Initially substantial amounts of *M. tuberculosis* containing particles are captured by mucus secreted in the airways and subsequently transported out of the body. Some particles may however bypass the body's first line of defence, and reach the alveoli of the lungs (Knechel, 2009). Alveolar macrophages rapidly surround and engulf the *Mycobacterium*, leading to phagocytosis of these bacteria (Willey, 2008).

Mycobacterium may continue to divide and multiply in these macrophages. In response to this, cytokines are produced in the host and subsequently T-lymphocytes accumulate (Knechel, 2009). The combination of macrophages and active T-lymphocytes creates a capsule around the bacterium, called a granuloma (Frieden *et al*., 2003). The formation of this granuloma, inhibit any further division and spread of the bacteria. In a healthy individual with a strong cell-mediated immunity, fibrosis and calcification of the lesion occur, leading to the successful control of the infection (Knechel, 2009). This stage is defined as the latent TB phase where the individual develops an immune response able to suppress the infection and keep the mycobacterium under control (Rueda *et al.,* 2010; Frieden *et al*., 2003). These individuals are asymptomatic, non-infectious and unable to transmit the disease.

The activation of latent TB to the active disease state is usually initiated when the individual's immune system becomes compromised, for whatever reason, enabling *M. tuberculosis* to emerge from the granuloma, resulting in reactivation of the infection (Tufariello, 2003). HIV is considered the leading cause for active TB progression; however, other conditions responsible for immunosuppression are uncontrolled diabetes mellitus, malnutrition, renal failure, chemotherapy, sepsis, smoking, organ transplantation and long term corticosteroid usage (Knechel, 2009). It is thought that only 5-10% of all individuals infected with *M. tuberculosis,* will actually develop active TB (WHO, 2010). TB generally occurs in the parenchyma of the mid and lower lung, known as pulmonary tuberculosis (Frieden *et al.,* 2003), causing symptoms that include coughing, chest pain, shortness of breath, night sweats, fever, fatigue, malaise and weight loss (Bakhsi, 2006). Although the lungs offer the most common place of infection, *M. tuberculosis* can spread to various locations in the body via the bloodstream and may cause clinical manifestations in the gastrointestinal and genitourinary tracts, bone, joints, nervous system, lymph nodes, skin and other organs. This condition is known as extrapulmonary tuberculosis (González-Martín *et al.,* 2010).

Although treatment with anti-TB drugs has shown to be effective, a vaccine that prevents TB infection and disease is needed for the control or elimination of TB worldwide (Flynn, 2004). The Bacille Calmette-Guerin (BCG) vaccine is known as the only TB vaccine available to date (González-Martín *et al.,* 2010). This vaccine contains an avirulent *M. bovis* strain and is administrated shortly after birth. The vaccine is considered effective in the prevention of TB among children, especially TB meningitis, however, it is shown to be less effective in adults (Flynn, 2004). Considering this, better vaccines are required to control TB.

2.3 Tuberculosis treatment

Most of the anti-TB drugs available today were discovered between 1950 and 1970, leading to a giant leap forward in the battle against TB. These anti-TB drugs, in combination with new therapeutic regimens, made it possible to cure this disease (Caminero *et al*., 2010). Anti-TB drugs, and their associated treatment regimens, are followed with the aim of curing the TB patient, to prevent death, to cease transmission, and prohibit the development of drug resistance (Frienden *et al*., 2003). Treatment regimes are considered suitable for use if 95% of all patients are cured from the disease and serious intolerance incidents are less than 5% (González-Martín *et al.,* 2010). These figures are however, less positive in many third world countries, where the picture is grim, and the chance of a successful treatment outcome is far less.

Treatment of the active TB disease usually involves the usage of a combination of anti-TB drugs with a variety of actions and functions (Graham, 2010). The anti-TB drugs may be divided into three groups: 1) bactericidal drugs, 2) sterilising drugs and 3) drugs that prevent drug resistance (Olivier & Loots, 2011).

Drugs classified as having bactericidal activity, are used to terminate and consequently prevent the replication of bacteria with active metabolisms, leading to a decrease in bacterial load, and prevent the transmission of TB. Isoniazid and rifampicin are generally considered the most effective bactericidal drugs used to date (Graham, 2010).

Rifampicin and pyrazinamide, are two important drugs with sterilising activity, which are used to eliminate *M. tuberculosis* with less active metabolisms, preventing clinical relapse. In order to prevent drug resistant *M. tuberculosis* from developing, bactericidal and sterilising drugs are usually used in combination, with the addition of one of the drugs, ethambutol or streptomycin (Graham, 2010). Anti-TB drugs may be classified as either first line or second line drugs.

First line drugs include the previously mentioned isoniazid, rifampicin, pyrazinamide, ethambutol and streptomycin. Second line drugs are used as alternatives for first line drugs, and administered only if the first line treatment fails. These second line drugs are usually considered less effective and may have more serious side effects (González-Martín *et al.,* 2010). These second line drugs include the fluoroquinolones, thioamides (ethionamide and protionamide), aminoglycosides (kanmycin and amikacin), polypeptides (capreomycin and viomycin), D-cycloserine and aminosalicylic acid (Caminero *et al*., 2010).

The WHO developed a strategy to control TB in 1995, known as the DOTS (directly observed treatment short-course) strategy, and has since been implemented in most countries worldwide (Atun *et al.,* 2005). The DOTS strategy consists of a six month treatment regime, where four first line anti-TB drugs (isoniazid, rifampicin, pyrazinamid and ethambutol) are used for the first two months, there after the patients are treated with isoniazid and rifampicin for the remaining four months (González-Martín *et al.,* 2010). The DOTS strategy also involves a health worker, appointed to ensure that each patient consumes the necessary medication, and completes the six months TB treatment regimen (González-Martín *et al.,* 2010). Health workers also document the sessions and provide additional counselling to patients (Bello, 2010). The patients used in my study followed the classical DOTS program as well (Hesseling *et al*., 2010).

As mentioned previously, a wide variety of anti-TB drugs are used in combination to treat TB. The obligated employment of multidrug regimens have, however, led to an increase in the associated anti-TB drug side effects (Gulbay *et al.,* 2006). The most common side effects are skin reactions, hepatotoxicity, gastrointestinal and neurological ailments. Hepatotoxicity occurs most frequently and is also considered the most severe of all the side effects (Tostmann *et al*., 2007). These side effects are also one of the main reasons for

discontinuation of anti-TB drugs and poor treatment compliance, resulting in treatment failure, treatment relapse or the development of drug resistance (Tostmann *et al*., 2007).

As previously mentioned, poor treatment compliance is one of the major causes for the increased prevalence of MDR-TB. MDR-TB can be defined as tuberculosis caused by *M. tuberculosis* strains that are resistant to both isoniazid and rifampicin (two most powerful first-line drugs) (Jain & Mondal, 2008). Extensively drug resistant (XDR) *M. tuberculosis* strains have also recently emerged, and are characterised by the resistance of isoniazid or rifampicin, any fluoroquinolone drug, and one injectable second line drug (Caminero *et al., 2010*). As previously mentioned, second line anti-TB drugs are consequently used to treat these cases. These anti-TB drugs are, however, as previously mentioned, less effective, more toxic and treatment takes up to 24 months (Frieden *et al*., 2003, Caminero *et al., 2010*). XDR tuberculosis, is however, even more difficult to treat and in some cases, surgery is necessary (Caminero *et al., 2010*). Furthermore, in 2007, the first totally drug-resistant TB (TDR-TB) cases were documented, showing resistance to all first-line and second-line anti-TB drugs (Velayati *et al.*, 2009). In a study done by Klopper *et al*. (2013) it was found that XDR and TDR-TB is spreading in South Africa, with most cases detected in KwaZulu-Natal, Western Cape and Eastern Cape provinces.

2.3.1 First line anti-TB drugs

As the aim of this study is elucidating the underlining mechanisms associated treatment failure to first line anti-TB medications, and potentially their side effects, this literature study will focus on these drugs, describing their mechanisms of action, in addition to their associated side effects.

2.3.1.1 Isoniazid

In humans, isoniazid (INH) is largely metabolised in the liver by means of two processes, 1) acetylation, by the enzyme N-acetyl transferase 2 (NAT 2), into acetylisoniazid and 2) hydroxylation into isonicotinic acid and acetylhydrazine (Huq, 2006). A number of studies have established that the rate at which isoniazid is metabolised varies significantly from one individual to another, therefore individuals are classified as rapid or slow acetylators

depending on their INH acetylation capacity (Ellard *et al*., 1972). Acetylhyrazine can subsequently be acetylated to form diacetylhydrazine or hydrolysed to produce isonicotinic acid and hydrazine (Tostmann *et al*., 2008). Hydrazine may also be metabolised to produce nitrogen, pyrvate hyrazone and 1,4,5,6-tetrahyro-6-oxo-3-pyridazine (Huq, 2006). Recent studies suggest that hydrazine, is the causative agent of isoniazid induced hepatotoxicity (Tostmann *et al*., 2008).

INH is considered to be effective against TB, is inexpensive and has minor toxicity (Ormerod, 2004). INH also has a high bactericidal activity and is used for eliminating active growing *M. tuberculosis.* INH is also mycobacteria specific and has almost no activity on other bacteria (ANNON, 2008), and is classified as a prodrug as it is activated by *M. tuberculosis* catalsase-peroksidase (KatG), when nicotinamide adenine dinucleotide (NAD), manganese ions and oxygen are present (Basso & Santos, 2005).

Once activated by the *Mycobacterium*, several reactive oxygen species and reactive organic radicals are generated, which in turn interact with and inhibit various targets in *M. tuberculosis* (Zhang, 2005). Inhibition of cell wall mycolic acid synthesis pathway, is considered the primary target for these generated oxygen species (Zhang, 2005). The gene, *InhA*, which codes for the long chain enoyl acyl carrier protein reductase (InhA), and is involved in fatty acid biosynthesis, was shown to be the target of the isonicotinic acyl radical (produced when isoniazid is activated by KatG) (Broussy, 2005; Zhang, 2005). During this process, the isonictinic acyl radical reacts with NAD, to produce a covalent NAD-isoniazid (NAD-INH) adduct, which in turn binds to the NAD(H) recognition site of the InhA enzyme, resulting in inhibition of its function (Janin, 2007). Under normal conditions, InhA uses NADH as a hydrogen donor, in order to catalyse the reduction of the trans double bonds of the carbonyl group of the fatty acyl substrates (Delaine *et al*., 2010). Consequently when InhA is inhibited by the NAD-INH adduct, the Fatty Acid Synthase II system (FAS-II) is blocked, preventing fatty acid elongation essential for mycolic acid biosynthesis (Delaine *et al*., 2010).

Resistance to INH is seen to result from a variety of insertions, deletions and point mutations, located in several different genes (Ahmad & Mokaddas, 2010). Mutations located in the *kat*G gene (responsible for INH activation), and mutations both in the regulatory and coding region of the *inh*A gene, result in the majority of the resistance to INH in *M. tuberculosis* (Laurenzo & Mousa, 2011).

When INH is used alone to treat TB, adverse effects rarely occur in individuals which do not have a history of kidney failure or liver disease. However, when used in combination with other anti-TB drugs, a variety of adverse effects are regularly seen (Arbex *et al*., 2009). Nausea, vomiting, arthralgia, headache, skin rash and fever, are of the minor side effects associated with INH consumption, and the use of additional medication in order to relieve these symptoms often helps (ANNON, 2008; Arbex *et al*., 2009). More severe side effects include neurological and psychiatric manifestations, peripheral neuropathy, polyneuritis and hepatitis. Peripheral neuropathy is the most common of these, occurring in approximately 20% of all treated cases. The co-administration of pyridoxine, usually reduces the symptoms of both peripheral neuropathy and polyneuritis (Arbex *et al*., 2009). The prevalence rate of hepatitis in patients treated with INH alone is far less, than when patients are treated with INH - rifampicin combination therapy. The prevalence and severity of these side effects are however linked to age and the consumption of alcohol (Arbex *et al*., 2009; ANNON, 2008).

2.3.1.2 Rifampicin

Three types of rifamycins are available for the treatment of TB: rifampicin, rifabutin and rifapentine (Caminero *et al., 2010*). Rifampicin (RIF), a lipophilic rifamycin derivative, is considered, the most powerful sterilizing drug currently used for treating TB and is used for the purpose of eliminating active and dormant *M. tuberculosis* (Shi *et al*., 2007; Ahmad & Mokaddas, 2010). RIF is considered the most important of all anti-TB drugs currently used, as it reduces the treatment period of TB to six months, when used in combination with other anti-TB drugs (Arbex *et al*., 2009).

Once RIF is orally ingested, 85% of the drug is metabolized in the liver and small intestine. The drug is rapidly eliminated mainly in the bile and only 30% is excreted in urine (Arbex *et al*., 2009; Sousa *et al*., 2008). RIF is considered a powerful inducer of the hepatic cytochrome P 450 (CYP450) system and undergoes progressive enterohepatic circulation and deacetylation by hepatic microsomal enzymes to produce the active metabolite, 25-desacetyl-rifampicin. RIF can also be hydrolysed to produce 3-formyl rifampicin which is also excreted in the urine (Sousa *et al*., 2008). RIF increase the metabolism of numerous drugs used in combination, and has also been seen to affect the plasma levels of various antiretroviral drugs (McCance-Katz *et al*., 2011; Tostmann *et al*., 2008). RIF functions primarily by binding to the β-subunit of the DNA-dependent RNA polymerase of *M. tuberculosis,* which is encoded by the *rpoB* gene.

When bound, RIF blocks elongation of the RNA chain, which consequently leads to inhibition of protein synthesis (Janin, 2007; Ahmad & Mokaddas, 2010; Laurenzo & Mousa, 2011). However, it should be noted, that although the molecular target of RIF has been properly classified, the exact mechanism by which this interaction leads to the death of *M. tuberculosis* is still largely unknown (Karakousis, 2009).

Miss-sense mutations within an 81 base pair region of the *rpoB* gene, is reported to account for 90-95% of all RIF resistant cases, leading to ineffective binding of the drug. Monoresistance to RIF is uncommon, and consequently, RIF resistance is often used as an indicator of MDR-TB, since 85-90% of all RIF resistant strains are also reportedly INH resistant (Ahmad & Mokaddas, 2010). It has also noted, that RIF resistance is associated with more treatment failures and patient deaths than resistance to any other anti-TB drug (Laurenzo & Mousa, 2011).

Minor adverse effects associated with RIF include: nausea, anorexia, abdominal pain, flu-like symptoms, dyspnea and headaches. RIF treatment has also been reported to result in discoloration of body fluids, resulting in orange-coloured tears, urine and sweat (Arbex *et al*., 2009). Furthermore 5% of all patients receiving RIF, have reported increased serum levels of hepatic enzymes and bilirubin, however, these levels normalize without the need for the discontinuation of the treatment. The mechanism of RIF-induced hepatotoxicity is still largely unknown and no toxic metabolite has been identified as yet. Less common severe adverse effects include: thrombocytopenia, leukopenia, eosinophilia, haemolytic anemia and acute interstitial nephritis, but are rare and occur in less than 0.1% of all treated individuals (Arbex *et al*., 2009, Tostmann *et al*., 2008).

2.3.1.3 Pyrazinamide

Pyrazinamide (PZA) was first synthesised in 1936 and since 1952, been used widely in the treatment of TB (Arbex *et al*., 2009; Caminero *et al.,*2010). PZA is classified as an important sterilisation drug, thus eliminating dormant *M. tuberculosis* bacilli located within hosts macrophages and other sites of acute inflammation (Ahmad & Mokaddas, 2010; Arbex *et al*., 2009). These dormant bacteria are also considered the causative agents for clinical relapses (Graham, 2010). PZA has also been considered the most effective anti-TB drug for treatment of latent TB (Sheen *et al*., 2009).

PZA, a nicotinamide analog pro-drug, requires amide hydrolysis by the bacterial enzyme, pyrazinamidase (a nicotinamidase), encoded by the *pnc*A gene, into its active form, pyrazinoic acid (POA) (Laurenzo & Mousa, 2011). It is also known that POA can be further oxidised by xanthine oxidase to produce 5-hydroxypyrazinoic acid (Tostmann *et al*., 2008). A part of PZA can additionally be directly oxidized by xanthine oxidase to produce 5-hydroxypyrazinamide which may undergo further deamination to finally produce 5-hydroxypyrazinoic acid. In one minor route of the PZA metabolic pathway, pyrazinuric acid is produced by means of conjugation of pyrazinoic acid with glycine. All these metabolites produced from PZA are eventually eliminated in the urine (Wu & Tsai, 2007; Huq & Hossian, 2006). Various mutations in the *pnc*A gene of *M. tuberculosis* are the primary cause of pyrazinamide resistance. These mutations are distributed along the entire gene, with 120 different mutations resulting in pyrazinamide-resistance reported to date (Laurenzo & Mousa, 2011).

PZA is a drug orally administrated and metabolized primarily in the liver, with 70% of this excreted in urine, and 3% remaining (Arbex *et al*., 2009). By passive diffusion, PZA enters the *M. tuberculosis* cells, where it is converted to POA. High concentrations of protonated POA consequently accumulate in the bacterial cytoplasm, as a result of an inefficient efflux system. This results in a decrease in the intracellular pH which consequently leads to inactivation of fatty acid synthase I, and consequently inhibiting mycolic acid biosynthesis and ultimately cell wall structure (Arbex *et al*., 2009). Additionally, the accumulating POA, is also thought to de-energize the cell membrane by collapsing the proton motive force, disrupting transport of various vital components across the cell membrane (Zhang, 2005). As an example, incorporation of both methionine and uracil is thought to be inhibited by POA, due to the decrease in intracellular pH and lowering of the membrane potential, in turn consequently results in a reduction of intracellular protein and RNA synthesis (Basso & Santos, 2005) of *M. tuberculosis*.

Membrane potential also plays an essential role in adenosine triphosphate (ATP) synthesis by F_1F_0 ATPase, as POA reduces the membrane potential. ATP production in dormant *M. tuberculosis* may consequently also be inhibited, leading to cell death (Basso & Santos., 2005).

In 2011, Shi et al. identified RpsA (ribosomal protein S1), as the potential target of POA. The binding of POA to RpsA, inhibits trans-translation, a crucial process for freeing rare ribosomes in non-replicating organisms. This process is expendable during active growth conditions of *M. tuberculosis*, however, it becomes vital for the bacteria in the management of stalled ribosomes or damaged mRNA and other proteins, under stress conditions. It is believed that the inhibition of the trans-translation process by PZA, interferes with the survival of *M. tuberculosis* under no replicating conditions. The above mentioned thus clarify the capacity of pyrazinamide as a sterilizing drug for the elimination of dormant bacteria.

Minor adverse effects associated with PZA, includes gastrointestinal symptoms like nausea, vomiting and anorexia. Hyperuricemia and arthralgia (joint pain), have also been reported, and are thought to occur as POA, inhibiting the renal tubular secretion of uric acid. Treatment with nonsteroidal anti-antihistamines and aspirin, has been shown to be effective against the pain. Additional side effects, including exanthema and pruritus may also be treated with antihistamines. Photosensitivity dermatitis has also been reported (Arbex *et al*., 2009; Ormerod, 2004; Gulbay *et al*., 2006; González-Martín *et al.,* 2010).

Some of the major adverse effects associated with PZA treatment includes, rhabdomyolysis with myoglobinuria and kidney failure, however, these are rare and if they occur, treatment with PZA is usually discontinued. Acute arthritis is also rather common in patients with a history of gout and is treated with allopurinol. PZA is also considered the most hepatotoxic of all the first-line anti-TB drugs, and for this reason, the dose of PZA is adjusted to the weight of the patient (Arbex *et al*., 2009). The mechanism of pyrazinamide-induced toxicity is largely unknown and it is still unclear as to whether PZA, or its metabolites, are responsible for the associated toxicity (Tostmann *et al*., 2008).

2.3.1.4 Ethambutol

Ethambutol (EMB) is an anti-TB drug used in combination with INH, RIF and PZA. It is considered an alternative to streptomycin, due to the fact that *M. tuberculosis* is less likely to become resistant to EMB (Jaber *et al*., 2009; Ahmad & Mokaddas, 2010). EMB is classified as a bacteriostatic agent, with structural similarity to D-arabinose.

This drug has been shown to be extremely effective for the elimination of both intracellular and extracellular *M. tuberculosis* (Laurenzo & Mousa, 2011; Shi *et al*., 2007). Once ingested, 75-80% of the EMB is absorbed, and metabolised in the liver.
EMB is firstly oxidized by liver microsomes to produce an intermediated aldehyde, which is then converted to the dicarboxylic acid, 2,2-(ethylenediimino)-di-butyric acid (Peets & Buyske, 1964). Approximately 50-80% of the drug is excreted in urine, of which 8-15% is excreted as these metabolites (Arbex *et al*., 2009).

The mechanism of action of EMB is very complex and largely unknown (Jaber *et al*., 2009). EMB is thought to function by disrupting the biosynthesis of arabinogalactan, a major polysaccharide of the mycobacterial cell wall, by inhibiting the enzyme arabinoyl transferase (three homologous membrane associated enzymes encoded by the *emb*C-*emb*A-*emb*B genes) (Jaber *et al*., 2009), the proposed target of EMB (Laurenzo & Mousa, 2011). The inhibition of this enzyme by EMB, prevents the polymerization of arabinose into arabinogalactan, and the synthesis of the mycolyl-arabinogalactan-peptidoglycan complex, leading to a decrease in cell wall permeability (Arbex *et al*., 2009; Laurenzo & Mousa, 2011; Zhang, 2005). Mutations located in the *emb*B gene of *M. tuberculosis* are considered to be the most common genetic alterations conferring resistance to EMB (Ahmad & Mokaddas, 2010). Approximately 47-62%, of all *M. tuberculosis* strains resistant to EMB, contain the mutation of codon 306 of the *emb*B gene (Laurenzo & Mousa, 2011).

EMB is considered a drug that is generally well tolerated by patients, with adverse effects mainly being dose and time dependent. Adverse effects associated with EMB include, retrobulbar neuritis, affecting the central fibres of the optic nerve leading to blurred vision and a decrease in visual acuity. These effects are however more common in older patients with impaired liver function. These side effects are also reversible, if detected in the early phase of treatment, and in such instances, EMB is immediately discontinued (Arbex *et al*., 2009). In rare cases, peripheral neuritis has been reported to occur, and the administration of pyridoxine relieves these symptoms. Other side effects include, gastrointestinal symptoms, like nausea, vomiting, abdominal pain, hepatotoxicity, cardiovascular symptoms (myocarditis and pericarditis) and haematological symptoms (eosinophilia, neutropenia and thrombocytopenia). Hypersensitivity and hyperuricemia may also occur (Arbex *et al*., 2009; Chung *et al*., 2008).

2.4 Tuberculosis treatment failure

Although tuberculosis is considered as a curable disease and well developed anti-TB drug regimens exist, many patients still don't respond with a positive treatment outcome at the end of the six month prescribed anti-TB treatment regimen. This occurrence is known as TB treatment failure (Table 2.1), can be defined as a positive sputum smear or bacteriological culture result at five months or later after the initiation of anti-TB drug treatment programme (WHO, 2010). Failure to cure this disease, subsequently results in an increased risk of developing drug resistant TB (MDR-TB and XDR-TB), as well as the spread of this in communities, which in turn will result in an increase morbidity and mortality rate (Namukwaya *et al*., 2011).

TB treatment failure is thought to occur due a one or more of the following reasons, including: anti-TB drug resistance, non-adherence to treatment, malabsorption of orally administrated drugs (Bento *et al*., 2009), anti-TB drug associated side-effects causing a lack of compliance, poverty and poor life quality, limited availability of recourses (Jindal, 1997), physician mismanagement, lack of patient knowledge of disease and treatment, ineffective or inappropriate prescriptions of anti-TB drugs, substance (alcohol) abuse (Lamsal *et al*., 2009), diabetes mellitus, HIV (Namukwaya *et al*., 2011) and smoking (Tachfouti *et al*., 2011).

The above mentioned contributors of TB treatment failure will subsequently be discussed in greater detail below.

2.4.1 Non-adherence to treatment and resistance to anti-TB drugs

Non-adherence to anti-TB treatment is considered a major problem in the management of TB. Treatment regimens present various challenges for adherence, due to the long duration of anti-TB drug regimen, the adverse side effects associated with these drugs and patients' feeling better long before the treatment regimen is completed. Non-adherence has also been stated as one of the leading causes of treatment failure in developing countries (Amuha *et al*., 2009).

It is reported, that the rate of adherence to standard anti-TB treatment regimens in developing countries, is a mere 40% (Bagchi *et al*., 2010). Anti-TB drug induced adverse side effects, are one of the major contributors to poor treatment compliance, due to the severity of some of these side effects, which is accompanied by additional pain and other secondary symptoms (Tostmann *et al*., 2007). In most cases, these TB patients suffering from these side effects, would rather discontinue the use of anti-TB drugs and live with the symptoms of the TB disease state, which is in most cases more bearable than the adverse side effects of the drugs. The long duration period of treatment also makes it difficult to adhere to the anti-TB treatment regimens, and a large number of patients discontinued the use of these anti-TB drugs, in the continuous phase (4 month phase with just isoniazid and rifampicin) of the DOTS programme (Amuha *et al*., 2009). The reason is that patients start to feel better and most symptoms of the disease have been relieved during this phase of treatment, thus the patients believe that the disease is cured and they stop administration of the anti-TB drugs.

As mentioned previously, treatment failure may also be caused due to the organism causing the infection, being a drug resistant strain. Drug resistance may be developed to the previously described non-adherence, inadequate treatment regimens or reinfection with a new strain of *M. tuberculosis* during the treatment phase (Sonnenberg *et al*., 2000). Drug resistant TB usually emerges due to treatment mismanagement, and can be spread between individuals in a population in the same manner as a drug-sensitive TB strain. This problem may worsen with the development or infection of MDR-TB and XDR-TB, as first line anti-TB drugs are no longer effective, and this subsequently results in treatment failure. Second line anti-TB drugs need to be used by these patients, however, these drugs are known to be less effective, more toxic, have increased side effects and are more expensive (Advaryu & Vakharia, 2011).

Furthermore, factors such as alcohol consumption (Jaiswal *et al*., 2003), drug abuse and smoking (Tachfouti *et al*., 2011) have been implicated as contributors of non-adherence and result in an increased risk of TB treatment failure.

2.4.2 Poverty, poor life quality and limited availability of resources

Poverty and poor living standards have been strongly correlated to TB. It is well known, that overcrowding in impoverished areas, dramatically increases the risk of infection and the spread of the TB disease. A more active infection rate of *M. tuberculosis* has also been correlated to malnutrition, sub-optimal working conditions and limited access to the various TB diagnostic and treatment facilities (Lamsal *et al*., 2009). Furthermore, the inflexibility due to employment obligations, results in poor clinic attendance, and subsequently poor or non-compliance to the DOTS program (Kumar & Singh, 2002).

In a study conducted in reigns of England, Wales and Northern Ireland, by Antoine *et al*. (2001), treatment outcomes of patients clinically diagnosed with TB were monitored. They found that 77% of the TB patients completed the prescribed treatment regimen of 12 months and only 16.5% of these patients were cured. They concluded that people living in urban parts of these reigns hand a higher percentage rate of treatment completion than those living in the rural areas (Antoine *et al*., 2001). All these mentioned factors associated with people living in poverty are major contributors to poor treatment compliance and eventually treatment failure.

2.4.3 Lack of patient knowledge and physician mismanagement

The illiteracy of many of the TB patients and their difficulty in understanding the information supplied to them regarding TB disease, effects and treatment, as provided by the health care workers, have also been implicated to be a contributor to treatment non-adherence and eventually treatment failure (Amuha *et al*., 2009). Furthermore, the cultural beliefs of these populations regarding the spread of TB, in addition to the misinformation given by the traditional healers, is still a major problem. In order to deal with the lack of TB knowledge among these infected populations, it was suggested that an increase level of TB disease and treatment awareness be implemented throughout the entire community, using teaching methods suitable for the population, in their language of preference (Lamsal *et al*., 2009).

Ineffective or inappropriate prescriptions of anti-TB drugs by physicians, are also commonly encountered outside the DOTS programme, and are also directly linked to TB treatment failure (Lamsal *et al*., 2009). A study done on 81 rural and 96 urban private medical practitioners in India, found that 49.6% of these practitioners used unacceptable treatment regimens, 30.1% of which was prescribed for an incorrect duration, and 19.5% prescribed regimens of poor quality (Uplekar *et al*., 1996). Reasons for this mismanagement of anti-TB treatment regimens may be attributed to 1) unqualified physicians, 2) lack of knowledge and awareness, 3) pressures and incentives offered by certain anti-TB drug brands, 4) patients' inability to afford the costs and 5) non-availability of drugs in some communities (Jindal, 1997).

2.4.4 Other disease states and contributing factors

Diseases which include Diabetes Mellitus (DM) and HIV have been reported to have a major impact on the treatment outcome of TB patients. In a study conducted by Baker *et al*. (2011), the impact of DM on TB treatment outcome was investigated. They concluded that DM increased the risk of anti-TB treatment failure, as well as death in these patients. Increased mortality and TB treatment failure have also been found in patients co-infected with HIV (Perriens *et al*., 1991). Another factor contributing to anti-TB treatment failure and associated with TB-HIV co-infection, is drug-drug interactions. The anti-TB drugs, isoniazid and rifampicin, are known to have inauspicious drug interactions with variety antiretroviral drugs used for HIV/AIDS treatment (Adhvaryu & Vakharia, 2011). These so-called drug-drug interactions, is thought to induce a number of side effects, including a decreased efficacy of the drugs, leading to treatment failure and subsequently to drug resistance, and also an increase risk of toxicity, and hence, treatment interruptions (Dean *et al.*, 2002).

Furthermore, it is known that isoniazid is able to permanently inhibit several CYP450 enzymes, until new CYP450 enzymes are synthesised. Rifamycin, on the other hand, induces several CYP450 enzymes, consequently leading to reduced therapeutic effects of numerous medications. This results into complicated health consequences when rifamycins are taken in combination with other well-known drugs, such as oral contraceptives, warfarin (Coumadin) or sulfonylureas (King & Bucker, 2009).

Table 2.1: Official WHO definitions of TB treatment outcomes

TB treatment outcomes	Definitions
Cured	A patient who was initially smear-positive and who was smear-negative in the last month of treatment and on at least one previous occasion
Treatment completed	A patient who completed treatment but did not meet the criteria for cure or failure. This applies to pulmonary smear-positive and smear-negative patients and to patients with extrapulmonary disease
Treatment failure	A patient who was initially smear-positive and who remained smear-positive at 5 months or later during treatment
Treatment defaulted	A patient whose treatment was interrupted for 2 consecutive months or more
Relapse	A patient previously declared cured but with a new episode of bacteriologically positive (sputum smear or culture) tuberculosis

2.5 Tuberculosis diagnostics

Tests and techniques used for TB diagnostics vary in speed, sensitivity, specificity and cost (Frieden *et al*., 2003). At present a wide variety of diagnostic methods are used in both cases of *M. tuberculosis* infection and TB disease, playing a pivotal role in the control of TB. The diagnosis of TB is, however, more difficult in older individuals, young children and immuno-compromised patients (Wolinsky,1994). Early diagnosis and effective anti-TB treatment, are the main objectives of TB control, and are aimed at reducing mortality and morbidity, and preventing the development of anti-TB drug resistance. To this end, monitoring the number of successful treatment outcome cases is an important indicator to assess the effectiveness of TB control programmes (Antoine *et al.*, 2006). In 2007, globally 270 000 (5%) of all patients diagnosed and treated for TB, relapsed within two years after treatment completion (Horne *et al*., 2010).

Considering this, monitoring treatment progression, and having the ability to accurately predict treatment failure well before the end of the 6 month treatment regimen, would be considered a major breakthrough, as non-responders could be identified early, and alternative measures taken, which would ultimately reduce the incidence of drug resistant TB (Horne *et al.*, 2010). Unfortunately, the accurate identification of high-risk patients to failure and relapse, does not yet exist, and hence, results in extended therapy, which is

problematic, especially in high-burden areas where poverty prevails (Menzies *et al.*, 2009). In the section to follow, the various TB diagnostic methods will be discussed, including their advantages and limitations, additionally considering their application to predicting treatment outcome.

2.5.1Techniques for the diagnosis of *M. tuberculosis* infection

2.5.1.1 Tuberculin skin test

The tuberculin skin test (TST) is considered the standard diagnostic test for determining tuberculosis infection since the early 1900s (Frienden *et al.,* 2003). Purified protein derivative (PPD), also known as tuberculin, is used, and consists of proteins obtained from *M. tuberculosis,* the BCG vaccine bacillus and other environmental mycobacterium (González-Martín *et al.,* 2010). TST is performed by the intradermal injection of PPD under the individual's forearm skin, and initiates a hypersensitivity reaction if *M. tuberculosis* infection is present. The reaction can be visualised as a thickening of the skin at the point of injection after a period of 24 to 48 hours (WHO, 2006).

This test, is however, not without its limitations. It has a low sensitivity and specificity, and is not able to distinguish between different mycobacterium species (Knechel, 2009). False-negative results may also occur in people who are immuno-compromised and who cannot initiate an immune response after injection. False-positive results may also occur in people that have previously received a bacillus calmette-guerin (BCG) vaccine or been infected with a different strain of mycobacterium.

2.5.1.2 Interferon Gamma Release Assays

Interferon gamma (IFN-γ) plays a critical role in regulating cell-mediated immune responses to *M. tuberculosis* infection. This led to the development of interferon gamma release assays (IGRAs) to detect *M. tuberculosis* infection by measuring IFN-γ release in response to antigens for *M. tuberculosis* (Pai *et al.*, 2006 a; Mazurek *et al.*, 2010). When *M. tuberculosis* is present, an increase of IFN-γ is noted in the blood, which indicates TB

infection (Fatima, 2009). Early versions of IGRAs were replaced by those using alternative antigens, more specific to *M. tuberculosis* than PPD. These antigens include early secreted antigenic target (ESAT-6), culture filtrate protein (CFP-10), and TB7.7 (Rv2654) (Pai *et al.*, 2006 a). These tests are based on the principle of an enzyme-linked immunosorbent assay (ELISA) to assess an individual's acquired immunological response to ESAT-6 and CFP-10, based on the amount of IFN-γ released (González-Martín *et al.*, 2010; Mazurek *et al.*, 2010).

The TST and IGRAs both have operational advantages, however IGRAs is more specific than the TST (González-Martín *et al.*, 2010). The advantages of IGRAs compared to TSTs, are that it has the ability to get results within a single visit, does not require two-step testing, and includes technical, logistic, and economic advantages (González-Martín *et al.*, 2010). However, this approach is not without its disadvantages, which includes a greater risk of test conversion due to false-positive IGRA results with follow-up testing of low-risk health-care workers who have tested negative at prior screening. False positive IGRA results are frequently not a true false positive result, however this technique is still more sensitive than the TST test and thus a person with a negative TST can have a positive IGRA (Mazurek *et al.*, 2001). In the light of the current investigation, when used for predicting treatment outcome or monitoring treatment progression, IGRAs was shown to be the more accurate indicator of progression to active TB over time than the TST, however IGRAs are more expensive (Kik *et al.*, 2010).

2.5.2 Active TB diagnostic Techniques

2.5.2.1 Chest radiology

Chest radiology is used for the detection of TB lesions and abnormalities in the upper and middle sections of the lungs, however, the disease activity cannot be establish with confidence (Konstantinos, 2010). Although this method is fast and convenient, chest radiology is also non-specific, due to the fact that a range of diseases can present similar chest radiographs.

Additionally, abnormalities on chest radiographs may indicate the possibility of TB, however, it is unreliable and cannot be used as a diagnostic method to confirm TB (WHO, 2006).

2.5.2.2 Microscopy techniques

Robert Koch initially demonstrated the first microscopic staining method for *M. tuberculosis* in 1882. Although microscopic advances occurred over the years, sputum smear microscopy remains the most used diagnostic method for the detection of TB (Drobniewski *et al*., 2003; WHO, 2006). The principle of sputum smear microscopy is to detect presence of acid-fast bacteria. Sputum samples obtained from patients are smeared to a glass slide, stained with a dye (basic fuchsin), dried, and subsequenty acid alcohol is then added. If acid-fast bacteria are present, the cells will remain red (Knechel, 2009). This method is, however, non-specific for *M. tuberculosis*, as other acid fast bacteria may also react in the same manner (Knechel, 2009).

Although sputum smear microscopy is rapid, inexpensive and easy to perform, it has limited sensitivity. Due to a poor detection limit (5000-10000 cells/ ml sputum), and hence, smear microscopy can only diagnose samples with high bacterial loads. This method is consequently, unreliable in especially children, people with HIV and cases of extrapulmonary TB, where bacterial loads are relatively low (WHO, 2009).

Treatment progression using this method can be followed, as the bacteria will visibly decrease with each smear – collected at different intervals after treatment has commenced. To date, several studies reported that a positive sputum microscopy during the second month of treatment is an indicator of treatment failure (Bernabe-Ortiz *et al.*, 2011).

2.5.2.3 Bacteriological cultures

TB diagnostics using culturing is considered the diagnostic gold standard. This method is shown to identify the *M. tuberculosis* bacterium in more than 80% of TB positive cases, with 98% specificity (Arslan *et al.,* 2010). Culture methods are more sensitive than smear microscopy techniques, with a detection limit of 10-100 bacteria/ml sputum (González-Martín *et al.,* 2010). Either solid or liquid media may be used, including: Lowenstein-Jensen, Kirchner and various Middlebrook formulations (Gray, 2004). *M. tuberculosis* bacteria are slow growing, and for this reason, an incubation period of 2-6 weeks is required, to ensure

sufficient growth and detection of the organism (González-Martín *et al.,* 2010). Species identification may follow in addition to antibiotic susceptibility tests (Gray, 2004).

The use of automated liquid media systems have improved on the long incubation period and made it possible to detect growth of the bacteria, within 7-10 days (González-Martín *et al.,* 2010). The first automated approach was the BACTEC 460TB system that utilizes a radiometric approach and media specific to mycobacterium. In this method, Middlebrook 7H9 medium, containing radiolabeled palmitic acid is used (Palomino *et al*., 2007). A mixture of antibiotics may also be added, preventing contamination. This approach functions by incubating the medium with the collected sample, and if mycobacteria are present, it would metabolize the radiolabled palmitric acid, releasing radioactive CO_2 gas, which is consequently detected (Palomino *et al*., 2007). There are, however, numerous limitations associated with the BACTEC 460TB system, which subsequently led to the development of non-radiometric culture systems, such as the BACTEC 9000 and the BACTEC MGIT 960 systems (Gray, 2004).

Culturing is also the preferred method for following treatment progression (Antoine *et al.*, 2006). Patient samples are taken bimonthly, 15 and again 30 days after the treatment has started (González-Martín *et al.*, 2010). A positive treatment outcome may be indicated by a negative *M. tuberculosis* culture within two months of commencing treatment.

2.5.2.4 Nucleic acid amplification (NAA) techniques

These techniques are based on the use of molecular methods, for the enzymatic amplification of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), ensuring a fast detection of mycobacterium in clinical samples (Knechel, 2009).

The most frequently used techniques, are the polymerase chain reaction (PCR), strand displacement amplification (SDA) and transcription mediated amplification (TMA) methods (WHO, 2006). These methods were initially thought to be highly specific and sensitive for the detection of *M. tuberculosis*, with as little as 10 bacteria per ml sample needed for a positive result, however, low sensitivity under field conditions has been reported (WHO, 2006). Compared to smear microscopy and culture methods, amplification techniques are largely

more sensitive than smear microscopy, but are to some extent less sensitive than culture methods (WHO, 2009).

Amplification techniques are widely used to confirm sputum smear positive results and to distinguish *M. tuberculosis* from other species of mycobacterium (Knechel, 2009; Frienden *et al.,* 2003). Additionally, these techniques also show promise in the use of materials other than sputum, which include blood, bone marrow, cerebrospinal fluid, lymph and urine (WHO, 2006). Amplification techniques are however expensive, complex and show higher false–positive results under field conditions (WHO, 2006). Furthermore, the Centra for Disease Control and prevention (CDC) guidelines do not recommend the use of NAA tests to monitor treatment response since NAA can amplify DNA from dead bacilli (Pai *et al.*, 2003), and as it only detects the presence of *M. tuberculosis* without actually giving the amount of bacteria present in the sample. Thus, it would not be possible to detect a decrease of bacteria for a positive treatment outcome.

2.5.2.5 Serology

The serology immunological methods used for TB diagnostics, function by identifying specific antigens and antibodies in blood samples (WHO, 2006). Although these methods are rapid, easy to perform and inexpensive (Wanchu, 2005; Gray, 2004), they have serious limitations in the sense that they are unable to accurately distinguish between latent and active TB, and between *M. tuberculosis* and other species of mycobacterium (WHO, 2006). This method is also less sensitive in children, patients with extrapulmonary TB and individuals suffering from HIV (WHO, 2006), and would not be suitable for use in monitoring treatment progression.

2.5.2.6 Phage assay

The phage assay is based on the principle of using a mycobacteriophage (a virus infecting mycobacteria) to infect and detect the presence of *M. tuberculosis* in clinical samples and culture isolates (Palomino *et al*., 2007; Pia *et al*., 2005). The FAST-Plague TB, is a rapid commercial method used for TB diagnosis, especially in countries with limited diagnostic facilities (Gray, 2004). The FAST-Plague TB method provides results within 48 hours and

has high sensitivity in smear-positive samples; however, low sensitivity has been noted in smear-negative, culture confirmed specimens (WHO, 2006).

Considering its mode of action, the phage-based assay could potentially be used as a fairly handy tool to follow treatment progression, especially in collaboration with other techniques. However, a major disadvantage may be that phage-based assays require approximately five times more volume of the sample compared with direct microscopy and culture (Palomino, 2005).

2.5.2.7 Xpert MTB/RIF assay

A more recently developed TB diagnostic technique, the Xpert MTB/RIF assay, endorsed by the WHO in Desember 2010, enables the simultaneous detection of *Mycobacterium tuberculosis* and rifampicin resistance (WHO, 2010). The Xpert MTB/RIF assay was specifically recommended to be used as an initial diagnostic test for suspected drug-resistant or HIV-associated pulmonary tuberculosis, and by June, 2012, two thirds of the high TB burden countries and half of the countries with a high MDR-TB burden, had incorporated this test into their national TB programme guidelines (Lawn *et al*., 2013).

The Xpert MTB/RIF assay is an automated real-time polymerase chain reaction assay, based on the amplification of a *M. tuberculosis* complex specific region of the *rpoB* gene, which is probed with molecular beacons to detect the presence of RIF resistance determining mutations (Theron *et al*., 2011). This assay is considered as highly sensitive and specific for the detection of TB and results are also available within two hours (Sekadde *et al*., 2013). In a study done by Nicol *et al* (2011), they found an overall sensitivity of 100% for smear positive - culture positive cases, 61,1% for smear negative culture positive, and finally a specificity of 98.8% when two induced sputum samples was analysed using the Xpert MTB/RIF assay, among children with suspected pulmonary TB in Cape town, South Africa (Nicol *et al*., 2011). Although the Xpert MTB/RIF assay is an accurate rapid rule-in diagnostic test for pulmonary TB and proven to outperform smear microscopy, it is however an expensive test to perform (Theron *et al*., 2011).

2.6 Predicting Treatment Outcome

Early diagnosis and effective anti-TB treatment, are the main objectives of TB control, and are aimed at reducing mortality and morbidity, and preventing the development of anti-TB drug resistance. To this end, monitoring the number of successful treatment outcome cases is an important indicator to assess the effectiveness of TB control programmes (Antoine *et al.*, 2006).

Monitoring treatment progression, and having the ability to accurately predict treatment failure well before the end of the 6 month treatment regimen, would be considered a major breakthrough, as non-responders could be identified early, and alternative measures taken, which would ultimately reduce the incidence of drug resistant TB (Horne *et al.*, 2010). Unfortunately, the accurate identification of high-risk patients to failure and relapse, does not yet exist, and hence, results in extended therapy, which is problematic, especially in high-burden areas where poverty prevails (Menzies *et al.*, 2009).

2.6.1 Predicting treatment outcome by sputum smear results

Until very recently, the WHO and the International Union Against Tuberculosis and Lung Disease (IUATLD) treatment guidelines, recommended sputum smear examination at the end of the initial phase (after 2 months), as a means of monitoring treatment progression. They suggested continuing with the initial phase for another month if the collected patient sputum smear is still positive at this stage, before the continuation phase is initiated (Feng-Zeng *et al.*, 1997; Horne *et al.*, 2010). However, in recent studies, conducted by means of a systematic review and meta-analysis, it was determined that this approach showed poor sensitivity and moderate specificity for predicting treatment failure or relapse (Horne *et al.*, 2010).

2.6.2 Predicting treatment outcome by anti-TB drug resistance

Treatment failure and relapse is low in patients with drug-susceptible TB receiving DOTS treatment, with a rifampicin-containing regimen (Weis *et al.*, 1994).

In these cases, the treatment failure rate is estimated at 1-4%, and relapse is at most 7% (Horne *et al.*, 2010).

Numerous studies have concluded that the administration of rifampicin-containing regimens, only in the initial phase of treatment, rather than throughout the whole course, was strongly associated with higher rates of treatment failure, relapse, and acquired drug resistance. Thus, these studies found that failure and relapse rates progressively declined with long duration of rifampicin, leading to a better treatment outcome (Menzies *et al.*, 2009). As these aspects are very important determinants of treatment outcome, monitoring or testing for rifampicin resistance in the early phase after initiating the treatment program, may potentially be used for predicting a poor treatment outcome.

Menzies *et al.* (2009) additionally indicated an important association of primary isoniazid resistance, with treatment failure, relapse, and acquired drug resistance. Thus, a widespread availability of rapid, inexpensive testing for isoniazid resistance is needed to monitor treatment outcome (Menzies *et al.*, 2009).

The methods to detect molecular resistance, are based on expression of mutations in specific genes. However, in some extracts they are not observed, and therefore undetectable (Ramaswamy & Musser, 1998). High specificity is achieved with screening protocols based on the simultaneous analysis of the mutant region and its corresponding wild sequence (González-Martín *et al.*, 2010). Mono-resistance to rifamycins is uncommon, and is often an indicator of MDR-TB, since most rifamycin resistant strains are also resistant to isoniazid (Ahmad & Mokaddas, 2010).

2.6.3 Predicting treatment outcome by monitoring bodyweight variation

Since TB is a progressive disease, bodyweight variation may be a practical anthropometric marker to predict TB treatment outcome. A study found that no significant difference exists between the two groups at baseline. However, during the first month of treatment, changes in weight over time amongst patients could be clearly seen – individuals with an adverse outcome, would lose an average of 0.97 kg, as opposed to those who were cured, who

showed an average increase of about 0.93 kg in weight. Furthermore, they found that weight gain in both outcome groups increased at the same rate after the first month of treatment. However, future studies with standardized measurements are needed to corroborate these findings (Bernabe-Ortiz *et al.*, 2011), as weight gain or weight loss may be associated with a number of disease and dietary conditions, and is not necessarily a specific marker for treatment outcome.

Considering all the evidence, it is clear that there is still great need for distinctive, more sensitive and specific biomarkers of a TB treatment response, and those more recently identified, have not yet been adequately validated in human trials, and until then, cannot be used as surrogate endpoints of treatment failure and relapse (Horne *et al.*, 2010). Furthermore, newer research approaches, such as metabolomics, may be used to identify new markers to monitor treatment outcome and/or the early prediction of treatment failure.

2.7 Patient collected sample materials used for characterising and diagnosing TB

2.7.1 Blood

Blood is considered one of the most frequent materials used in general TB research, due to the fact that blood samples are easy to collect, involves minimum amount of pain and therefore preferred over sputum samples which is often difficult to obtain and where invasive techniques are needed (Cruz *et al*., 2011). In TB diagnostic methods, which include culturing and nucleic acid amplification techniques, blood is commonly used for diagnosis of extrapulmonary TB (Green *et al*., 2009).

In techniques used for diagnosis of TB infection, blood is also used. The interferon- γ assays (IFN-γ) is one of these techniques, where the amount of interferon- γ is measured in the blood. When *M. tuberculosis* is present, an increase of interferon- γ is noted in the blood, which indicates TB infection (Fatima, 2009). Blood is also used as material for the measurement of immunological markers, which may be used as indicators of disease progression or treatment response in patients suffering from TB (Perrin *et al*., 2007). Whole-blood bactericidal assay is another example where blood is used. This is a model of

intracellular infection and has been proposed as a biomarker of sterilising activity of anti-TB drugs, used to assess new anti-TB drug candidates (Parida & Kaufmann, 2010; Perrin *et al*., 2007). In a recent study conducted by Cruz *et al.,* 2011, a Nested-PCR (NPCR) method was used for the detection of *M .tuberculosis* in blood and urine samples obtained from patients suffering from TB. They detected, when using NPCR on blood samples, 72.4% accuracy is obtained, in the diagnosis of TB confirmed patients, however, the sensitivity of PCR is known to vary (Cruz *et al.,* 2011). It is also said that haemoglobin in blood can be considered as an inhibitory substance, which could interfere with these PCR methods (Cruz *et al.,* 2011).

In another study conducted by Tanaka *et al.,* 2011, they used a proteomic approach in whole blood supernatants, in an attempt to find candidate biomarkers, which may possibly discriminate cured TB patients, from those with a risk of relapse. Three TB-associated proteins, retinol binding protein 4 (RBP4), Fetiun-A, and culsterin, were identified in whole blood supernatants. They found that the levels of these proteins were significantly reduced in patients diagnosed with active TB (Tanaka *et al.,* 2011).

2.7.2 Sputum

Sputum can be defined as a thick, mucus like substance, which is coughed up from the lungs and has been widely used for general TB research for many years. Microscopic examination of stained sputum is considered the most common used method in TB diagnostics (Arslan *et al*., 2010). This method has, however, low sensitivity, since 5000 to 10 000 *M. tuberculosis* cells must be present in one millilitre of sputum (WHO, 2006). Different *Mycobacterium* species cannot be distinguished and the technique has no value in cases of HIV co-infected and extrapulmonary TB (Perrin *et al*., 2007). The diagnostic gold standard of TB diagnostics, bacterial culturing also uses sputum as clinical material. This technique has high sensitivity and can detect as little as 100 *M. tuberculosis* cells per one millilitre of sputum (Palomino, 2005). Sputum culturing is, however, time consuming and contamination may compromise results (Perrin *et al*., 2007). Sputum is considered a good substrate for PCR analysis (Cruz *et al.,* 2011), and is used in various nucleic acid amplification techniques for the detection of *M. tuberculosis* (Green *et al*., 2009), however, sputum also contains inhibitory substances, like polysaccharidesl, which could interfere with PCR reactions (Cruz *et al.,* 2011). Sputum samples have also been used in a number of metabolomic

applications for the measurements of biomarkers in disease conditions (Parida & Kaufmann, 2010; Schoeman, du Preez & Loots, 2012; Olivier & Loots, 2012).

Sputum is however, not the ideal material to use for general research, since 1) children suffering from TB are unable to produce sputum for analysis (Cruz *et al.,* 2011), 2) in some cases patients who have improved upon anti-TB treatment, stop expectorating sputum too early (Perrin *et al*., 2007), and 3) the use of sputum samples obtained from patients with HIV-associated or extrapulmonary TB, is not very relevant (Green *et al*., 2009). Collecting sputum samples from patients is not considered an invasive procedure, however it is seen as being more invasive than collecting a urine sample.

2.7.3 Urine

Urine has been considered an imperative source for markers of health and disease, as early as 4000 BC, when Hippocrates thought urine was a filtrate of the humors (historical reference to the fluids of the human body) and used the urine as a diagnostic tool for diseases. In the present, urine analysis has been strongly established as a tool for disease diagnoses and involves physical, chemical and microscopic analysis (Ryan *et al.,* 2011). The use of urine in comparison with other biofluids, as a tool for general research, has a number of advantages, which include: 1) large quantities can be obtained by means of a non-invasive manor, 2) repeated sampling is no problem, 3) less complex pre-treatment is needed, 4) urine samples are easy to collect and urine can be used for biomarker measurements in varies diseases including TB (Ryan *et al.,* 2011; Parida & Kaufmann*,* 2010; Olivier & Loots, 2012; Schoeman, du Preez & Loots, 2012).

In cases of the disease tuberculosis, urine is considered as an effective diagnostic material in both pulmonary and extrapulmonary TB, giving urine a major advantage over sputum, which is mainly limited to diagnosis of pulmonary TB (Gopinath *et al.,* 2009). Urine has been widely used in TB diagnostic methods such as culturing and nucleic acid amplification techniques. Most recently urine has been used for trans renal DNA diagnosis of TB, where quantitative PCR methods are used for the detection of trans renal mycobacterial DNA, thought to be present in urine. This technique, however, varies in sensitivity and specificity (Green *et al*., 2009).

In the past urine has traditionally been used to measure one or at the most a few metabolites. Over the years the need, of a holistic approach to metabolism led to metabolomics of urine for disease diagnosis (Ryan *et al.,* 2011). Various phenotypic, physiological and external factors influence the human metabolome, including gender, age, stress, medication, exercise and fasting, to mention only a few. These influences can potentially be distinguished by means of urine metabolomics (Ryan *et al.,* 2011). Metabolomics have also been widely used in the identification of biochemical markers in urine obtained from patients suffering from kidney cancer (Kotlowska *et al*., 2011), prostate cancer and gastric cancer (Wu *et al*., 2010). A major disadvantage in the use of urine for metabolomics studies is the large biological variations in urine composition (Ryan *et al.,* 2011). Creatinine is present in a number of body fluids including urine. The concentrations of endogenous metabolites may be expressed as ratios relative to creatinine, and for this reason creatinine, is used to normalise urine metabolites and to correct for effects like urine volume and concentration (Ryan *et al.,* 2011).

Table 2.2: Potential advantages and disadvantages of blood, sputum and urine as diagnostic and research material.

Sample type	Advantages	Disadvantages
Blood	• Can be used for biomarker measurements in TB cases (Perrin *et al.,* 2007). • Effective in pulmonary and extrapulmonary TB diagnosis (Cruz *et al*., 2011) • Effective substrate in PCR analysis for TB diagnosis (Cruz *et al*., 2011) • Can be used to identify TB-associated proteins using a proteomic approach (Tanaka *et al*., 2011)	• Invasive techniques are used to obtain blood • Pre-treatment of blood is needed • Contain inhibitory substance, haemoglobin, which interfere with PCR reactions (Cruz *et al*., 2011)
Sputum	• Can be used for biomarker measurements in TB cases (Parida & Kaufmam, 2010). • Most effective material used in the gold standard TB diagnostic method for pulmonary TB diagnosis. • Effective substrate in PCR analysis for TB diagnosis (Cruz *et al*., 2011)	• The use of sputum is insufficient for extrapulmonary TB diagnosis (Perrin *et al.,* 2010) • Complex pre-treatment of the sample is needed (Ryan *et al.,* 2011). • Sputum is less effective in TB diagnosis of HIV infected people and children (Green *et al*., 2009)

		• Contains inhibitory substances like, polysaccharides which interfere with PCR reactions (Cruz *et al*., 2011)
Urine	• Effective in both pulmonary and extrapulmonary TB diagnosis (Gopinath *et al.,* 2010). • Can be used for biomarker measurements in TB cases (Parida & Kaufmann, 2010). • Large quantities can be collected (Ryan *et al.,* 2011). • No hash or invasive techniques are used to obtain samples (Ryan *et al.,* 2011). • Less pre-treatment of the sample is needed (Ryan *et al.,* 2011).	• Urine concentrations vary but it can be corrected • Many factors interfere with normalization of urinary metabolites (Ryan *et al.,* 2011) • Standardization of urine can be complex (Ryan *et al.,* 2011)

2.8 Metabolomics

Metabolomics, is defined as the non-biased identification and quantification, of all the metabolites present in a biological system, using highly selective and sensitive analytical methods (Dunn *et al., 2005)*. Thus, metabolomics aims to give a global overview of the metabolome of an organism and plays a crucial part in the study of metabolites, produced in response to pathophysiological stimulations and genetic modifications (Kotlowska *et al.,* 2010).

It is well known that every organism has a complete set of small metabolites. When a disease like tuberculosis infiltrates an individual, the bacteria with their own unique set of metabolites, disturb the biological network of the host, and initiate changes in its metabolite profile (Parida & Kaufmann., 2010). Highly specialised methods can be used to measure the metabolite changes in response to these influences, potentially leading to the identification of new biomarkers, which may give more insight into disease mechanisms. A biomarker may be defined as a compound that may be used as a measurable indicator for a normal

biological or pathogenic state or a pharmacological response (Perrin *et al., 2007*). A combination of biomarkers may be termed, a biosignature (Parida & Kaufmann., 2010).

In metabolimics, a wide variety of highly sophisticated techniques are used that are able to simultaneously analyse, identify and quantify thousands of metabolites present in a biological sample (Kaddurah-Daouk *et al*., 2008). Gas chromatography (GC), liquid chromatography (LC), mass spectrometry (MS), capillary electrophoresis (CE) and nuclear magnetic resonance (NMR), are just some of the many instruments used in metabolomics. The use of combined technologies, such as GC together with MS, and LC with MS, is most preferably used in order to get the best possible results and reliable metabolite profiles (Kaddurah-Daouk *et al*., 2008, Oresic., 2009).

Over the years, metabolic signatures (biosignatures) have been identified for several disease conditions using metabolomics, including cancer, type 2 diabetes mellitus, cardiovascular disease, neurological disease, hypertension, preeclampsia, Huntington's disease, Alzheimer's disease, depression and schizophrenia (Madsen *et al., 2009,* Kaddurah-Daouk *et al*., 2008). Furthermore, metabolomics has also proven to present researchers and diagnostic groups with a number of advantages, as it is able to analyse the end products or metabolites present in relatively easily attainable biological specimens, including; blood, sputum and/or urine, which may have altered due to the disease state. Additionally, metabolomics also allows for the identification of metabolite markers that are both endogenous and exogenous to the organism, infecting the human host, also shedding light on bacteria-host interactions (Schoeman & Loots, 2001). Considering this, metabolomics may be used as a tool to differentiate patients clinically diagnosed with TB from healthy individuals on the basis of the variation in their metabolite profiles, for the purpose of identifying distinctive and specific metabolite markers, which better characterise the disease, and/or could potentially also be used diagnostically. These biomarkers should then normalise once treatment commences, and could potentially be used to predict treatment outcome (Wallis *et al*., 2009).

Furthermore, the identification of metabolic signatures for a range of drugs has also been made, shedding light on possible drug mechanisms and associated side effects (Kaddurah-Daouk *et al*., 2008, Huo *et al.,* 2009). Drug induced side effects have also been implicated

as major contributors of poor treatment compliance and eventually resulting in treatment failure. In a study conducted by Hewer *et al*., 2006, a H-NMR-based metabolomics approach was used to distinguish between three groups of patients including, HIV-1 positive/AIDS patients utilizing antiretroviral therapy, HIV-1 positive/AIDS patients not receiving antiretroviral therapy and HIV-1 negative individuals (Hewer *et al*., 2006). Separation of the three groups was achieved and a clear distinction could be made based on the metabolic profiles. They concluded that antiretroviral associated side effects could be monitored with the use of NMR-based metabolomics (Hewer *et al*., 2006).

2.8.1 TB Metabolomics Research to Date

Loots *et al*., 2004, also used a metabolomic research approach to study the possible effects of a combined anti-TB drug, Rifater (made up of, rifampicin, isoniazid and pyrazinamide), by studying the changes induced by the drug on the urine organic acid and free radical profiles of Sprague Dawley rats (Loots *et al*., 2004). The Rifater treatment increased the hydroxyl radicals and various organic acids characteristic of a multiple acyl-CoA dehydrogenase defect (MADD), and subsequently suggested that Rifater treatment induces its associated side effects by inhibition of either the electron transport flavoprotein (ETF) or the ETF dehyrogenase enzyme. They reported that the co-administration of melatonin reduced the levels of the hydroxyl radicals and restored the abnormal organic acids profile caused by Rifater (Loots *et al*., 2004). These studies also prove the capability of metabolomics as a tool for the elucidation of drug action and associated side effects.

Since then, the Infectious Disease Metabolomics research group at the North-West University, Potchefstroom, South Africa, headed by Prof Du Toit Loots, used this approach for answering a number of TB related questions. Olivier & Loots (2012), identified markers better characterising various TB causing and non-TB causing *Mycobacterium* species. With the use of GC-MS lipid analysis on sputum, in combination with multivariate statistical data processing, they constructed a model consisting of 12 of these metabolite markers, best describing the variation between the various *Mycobacterium* species, for the purpose of applications to diagnostics and subsequently patented this model, and published a number of papers on the topic (du Preez & Loots, 2012; du Preez & Loots, 2013; du Preez & Loots, 2013).

In the same year, Schoeman, du Preez & Loots (2012) also compared four sputum pre-extraction preparation methods for the identification and characterisation of *Mycobacterium tuberculosis*, using GCxGC-TOFMS metabolomics. The aim of this study was to find the best sputum pre-extraction preparation method, due to the fact that reproducible/repeatable results during compound extraction and analysis are difficult to obtain due to sputum's high viscosity and uneven consistency. The four sputum pre-extraction preparation methods were compared using: 1) Sputolysin; 2) a combination of N-acetyl-l-cysteine and sodium hydroxide; 3) sodium hydroxide alone, and 4) a simple ethanol homogenisation method. They found that the last method, a simple ethanol homogenisation approach, proved to be superior. Using this sputum pre-extraction preparation method, they subsequently differentiated between *M. tuberculosis* spiked sputum and control sputum sample groups, at a detection limit of only 100 cells, using as little as 250 µL sputum. Additionally they were able to clinically diagnose TB-positive and TB-negative sputum sample groups at a specificity and sensitivity equally, or perhaps even better than that of bacteriological culture. They concluded that this method has the capacity to identify metabolite markers indicative of the physical presence of *M. tuberculosis* in patient collected sputum, as well as metabolites which are characteristic of an altered TB induced host metabolism, which is extremely useful for investigating the impact of the TB disease state on the human metabolome. They further stated that this approach may also be applied to characterise various other lung disease, using sputum (Schoeman, du Preez & Loots, 2012).

Considering this, it is clear that metabolomics could contribute significantly in achieving the aims set out in this study and could provide much needed explanations for better characterising why certain individuals fail to respond to treatment, via the identification of new metabolite markers associated with treatment outcome. Furthermore, the identification of such markers may also potentially be applied diagnostically, in order to monitor or predict treatment outcome, well before the 6 month treatment regimen is completed.

Chapter 3

Aims and Objectives

3.1 Problem statement

The administration of anti-TB drugs and their associated treatment regimens are designed with the aim of curing the TB positive patient. However, in many instances, diseased individuals don't respond with a favourable treatment outcome, even after completion of the entire 6 month anti-TB treatment regimen. This phenomenon, known as TB treatment failure, is still poorly understood and remains a major issue globally. As previously mentioned, numerous factors are implicated as major contributors of TB treatment failure, which include a long duration period of anti-TB drug administration, non-adherence to anti-TB treatment, adverse side effects associated with anti-TB drugs, poverty and poor life quality, multidrug-resistant TB and lack of development of new and more effective anti-TB drugs. One aspect of this, that isn't clearly addressed in the literature, is the underlying variation in each patient, resulting in his/her varying reaction to the anti-TB drug regimen, and hence its varying efficacy from one patient to the next. Furthermore, little is known about the underlying variation of the host to the primary TB infection or response to the TB disease state, and how some patients have more effective mechanisms for eliminating the infection, or recovering from the disease, either with or without treatment. Many factors may contribute to this, including diet, social practices (drinking and smoking), and perhaps even genetics. These mechanisms are poorly understood, and very little research has been done on treatment failure from this perspective. This metabolomics study, however, may identify the much needed metabolite markers, which may shed light on treatment failure, from this perspective, as it focuses on the response of the human metabolome to infection or disease, and the medication used for treating this, comparing those individuals with a positive treatment outcome (cured), to those in whom the treatment approach failed (treatment failure), comparing these groups at various stages during the treatment protocol.

3.2 Aims

1. Validation of the methodology used for metabolomics analysis of patient collected urine in the context of this investigation (refer to Objective 1).

2. To use the GC×GC-TOFMS metabolomics research approach validated in Aim 1, to identify potential biomarkers, better characterising or predicting treatment failure, by comparing the urinary metabolite profiles of those individuals successfully treated vs those where treatment failed, at different time points throughout the 6 month treatment regimen (Objectives 2-4).

3.3 Objectives

The aims will be completed through the following objectives:

1) Validation of the metabolomics methodology by investigating the 1) repeatability of the extraction procedure/analyst's ability for performing this, and 2) the GCxGC-TOFMS analytical repeatability for the compounds extracted using this approach, for the metabolomics research applications, in the context of this investigation.

2) To compare the GC×GC-TOFMS analysed urine organic acid profiles, of these TB positive patients at time of diagnosis, in order to identify potential biomarkers, better characterising the differences which may exist between those who were eventually cured and those where treatment failed, before commencing treatment, possibly explaining variation in these groups due to their varied response to the TB disease state.

3) To compare the GC×GC-TOFMS analysed urine organic acid profiles of these patients, during the 6 months DOTS treatment regimen, in order to identify potential biomarkers, better characterising the differences which may exist between those who were eventually cured and those where treatment failed, possibly explaining variation in these groups due to their varied response to the treatment regimen. In order to accomplish this, the difference in metabolite concentrations at week 4 of treatment and week zero (before treatment) were calculated, and univariate and multivariate comparisons between the cured and failed treatment outcome groups were done using these calculated differences.

4) To compare the GC×GC-TOFMS analysed urine organic acid profiles, of these patients, considering all time points simultaneously, during the 6 month treatment regimen, using a more conventional statistical approach, in order to identify new biomarkers or confirm the previously identified biomarkers, better characterising the differences which may exist between the those who were eventually cured and those where treatment failed, and the causes for these differences.

Chapter 4

Materials and Methods

4.1 Reagents

Ethylacetate and di-ethylether were purchased from Honeywell International Inc., (Muskegon, MI, USA). These organic solvents were ultra-pure Burdick & Jackson Brands. Hyrdochloric acid (HCl), pyridine, sodium sulfate anhydrous (Na_2SO_4) and trimethylchlorosilane (TMCS) were purchased from Merck (Darmstadt, Germany). 3-Phenyl butyric acid, hexane and O-bis(trimethylsily)-trifluoracetamid (BSTFA) were purchased from Sigma-Aldrich (St. Louis, MO, USA).

4.2 Urine samples

Urine samples from 46 patients clinically diagnosed with active TB were collected from baseline to the end of the six month classical DOTS treatment regimen, by the faculty of Health Sciences at the University of Stellenbosch. These patients are mainly from the rural areas located in the Western Cape, where the TB notification rate was determine as 841 per 100 000 population at time of sample collection. Patients were selected from five primary health care TB clinics within the public health system (Hesseling *et al.,* 2010). Sputum and urine samples were collected from patients suspected of having TB, based on a medical assessment of the symptoms associated with the active TB disease, by trained clinical staff, in sterile polypropylene screw-top vessels. Urine samples were immediately frozen at -80^{o}C and sputum samples were sent to a laboratory where standard diagnostic procedures, which included bacteriological cultures and Ziehl-Neelsen staining, were performed. A positive result on both diagnostic procedures confirmed the presence of the active TB disease in these 46 patients.

These patients immediately commenced treatment after a positive diagnosis for TB was made. The standard five days a week, clinical-based DOTS anti-TB treatment was given to these patients by routine clinic nurses. Fixed-dose combinations of anti-TB drugs were used, with dose adjustment based on patients body weight. Treatment consisted of a two month intensive treatment phase of the following anti-TB drugs: isoniazid (INH), rifampicin (RIF), pyrazinamide (PZA) and ethambutol (EMB) for two months, followed by a four month continuation phase with INH and RIF (Daily dosages for patients < 50 kg: RIF 480 mg, INH 320 mg, PZA 1000 mg and EMB 800 mg; and for those > 50 kg RIF 600 mg, INH 400 mg,

PZA 1250 mg and EMB 1200 mg)(Hesseling *et al.,* 2010). All the patients received the same treatment regimen. Urine samples were collected at 5 different time intervals, and included a sample at time of diagnosis and samples collected after week 1, 2, 4 and 26 of the treatment regimen.

The collected urine samples were requested from the DST/NRF Centre of Excellence for Biomedical TB research, from where they were transported to the North-West University (NWU) - Centre of Human Metabonomics, Infectious Disease Laboratory, on dry ice, for the metabolomics analyses pertaining to this study, and again stored at - 80 °C until further metabolomic analysis. Ethical approval for this study was obtained beforehand from the relevant Ethics Committee of the North-West University (No.NWU-00127-11-A1) and Stellenbosch University (No.99/039), as well as an informed consent from each participant for the use of their urine and sputum for research purposes. All samples were anonymised in order to maintain confidentiality. Clinical information accompanying these samples included; age, gender, body mass index (BMI) and treatment outcome. Microbial information, regarding results of the bacteriological cultures and Ziehl-Neelsen staining procedures, obtained from patients at time of diagnosis and week 26 of treatment period, was also provided (Appendix 1, Table 1).

The urine samples collected from the 46 patients were classified into two groups: 1) those from patients who responded positively to the anti-TB drug treatment (successful treatment outcome) i.e. those who showed a negative sputum culture result at week 26 (n = 29), and 2) those from patients who were not successfully treated (unsuccessful treatment outcome) and still showed a positive sputum culture result at week 26 (n = 17). Figure 4.1 summarises the urine sample collection over the different time points.

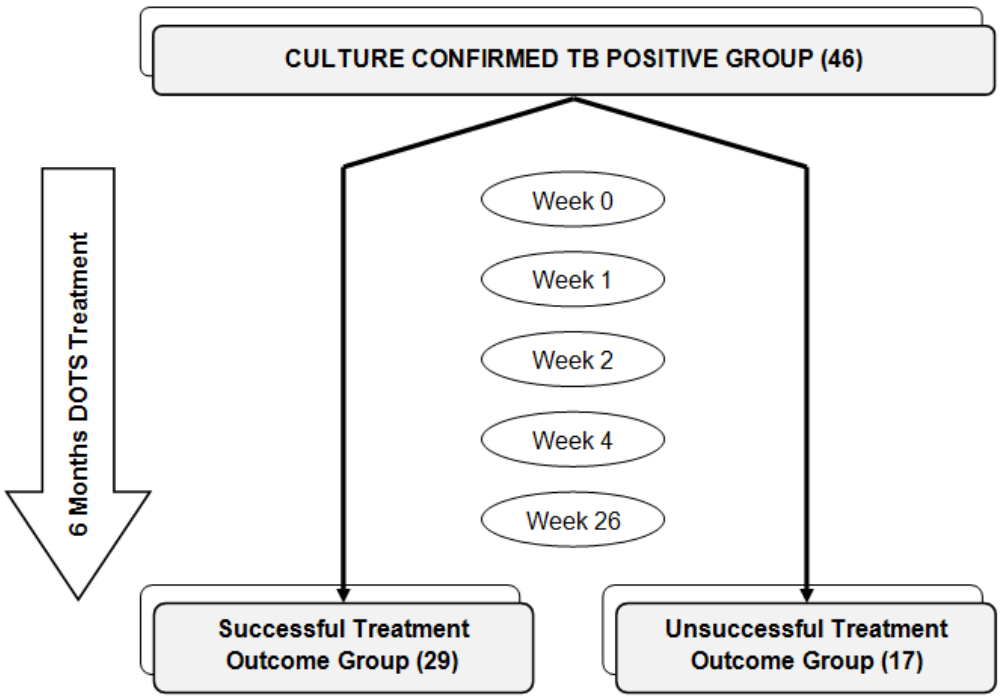


Figure 4.1 Summary of the urine sample collection time intervals.

In order to accomplish the aims of this study, all the collected urine samples, from all 46 patients, collected at all the time periods, from baseline to the end of treatment, were analysed using the metabolomics approach described below. Samples were analysed in batches of 26 urine samples, 23 of which were randomly selected patient urine samples, and 3 were quality control samples.

4.3 Quality control samples

Quality control (QC) samples were used to determine and correct for any inconstancies that may have occurred during the extraction or analysis of the collected urine, as well as to provide a representative mean sample containing all analytes which might be encountered during organic acid analysis. These QC samples were made up, by combining equal fractions (0.5 ml) from each sample in the entire patient sample cohort, and after mixing, aliquotting 1,5 ml of this stock QC urine sample mixture into 2 ml sterile tubes. These were subsequently frozen at – 20 °C until extraction and GCxGC-TOFMS analysis.

One QC aliquot sample was subsequently extracted per patient urine sample batch analysed, and injected three times, at the beginning, in the middle and at the end of each batch analysed. Prior to analysis of the first batch in the injection sequence, the respective QC was injected at least 6 times in order to equilibrate the analytical apparatus. QC samples were extracted, derivatized and analysed in the same manner as each patient urine sample collected, as described below (sections 4.4.2). The GCxGC/TOFMS repeatability as well as the repeatability of the organic acid extraction and derivatization procedure, was also determined using the data collected from these QC samples.

4.4 Methods

4.4.1 Creatinine determination

Creatinine is derived from creatine phosphate in muscle and is found in various body fluids such as urine, cerebrospinal fluid, gastrointestinal fluids, sweat, bile and serum. This metabolite is secreted in relatively constant amounts in an individual; however, individual creatinine values are dependent on the age, gender and lean body mass (Ryan *et al.,* 2011).

The creatinine value is used to compensate for varying urine concentrations, in order to ensure accurate comparison of the analysed metabolites between individual patients, by calculating the concentration of these metabolites relative to the creatinine value. Furthermore, the volume of urine to be used for the analysis, the volume of internal standard to be added, and also the amount of BSTFA, TMCS and pyridine required for derivatization (Loots *et al*., 2004), is also determined using this value (Tables 4.1 and 4.2). The creatinine value of each urine sample was determined before the extraction procedure by AMPATH Laboratories.

Table 4.1: Volume of urine used in the organic acid extraction method, according to creatinine values.

Creatinine value	Amount of urine
Creatinine value > 8.8 mol	Use 0.5 ml urine
Creatinine value < 8.8 mol, however > 0.44 mol	Use 1 ml urine
Creatinine value < 0.44 mol, however > 0.18 mol	Use 2 ml urine
Creatinine value < 0.18 mol	Use 3 ml urine

Table 4.2: Formulas used to determine the needed volume of reagents used in the organic acid extraction method.

Formula to determine urine volume	Reagents
5 × Creatinine mg% = volume in µl	Internal standard
3 × Creatinine mg% = volume in µl	BSTFA
0.6 × Creatinine mg% = volume in µl	TMCS
0.6 × Creatinine mg% = volume in µl	Pyridine

4.4.2 Extraction and derivatization of organic acids

The organic acid extraction method used in this study, is similar to that described in Loots *et al*. (2004).

Prior to organic acid analysis, urine samples were thawed and sonicated to ensure that all the precipitated material dissolved or became resuspended. An amount of urine, as determined from the creatinine value (section 4.4.1, Table 4.1) was transferred to a silanized glass tube (Kimax) and acidified by addition of 6 drops 5N HCl to adjust the pH of the urine to < 2. The internal standard (3-phenylbutyric acid) was added to a final concentration of 180 mmol/mol creatinine. 3-Penylbutyric acid was chosen as internal standard as it is absent in normal urine, elutes almost in the middle of an organic acid profile and is known to co-elute with very few organic acids (Reineke *et al*., 2012). Ethylacetate (6 mL) was then added to each urine sample and shaken on a rotary wheel for 30 minutes. After centrifugation at 3000 rpm for 3 minutes to facilitate phase separation, the organic phase (top phase) was removed

with a Pasteur pipette and placed into a clean Kimax test tube. Diethylether (3 mL) was subsequently added to the aqueous phase and the solution was shaken for a further 10 minutes. After a second centrifugation step, the upper organic phase was removed and added to the first organic phase (ethyl acetate phase). To ensure no residual water is present in the collected organic phase mixture, a small amount (approx. 2-5 g) of anhydrous sodium sulphate (Na_2SO_4) was added to the collected and combined organic phase mixtures. The tubes was briefly vortexed and subsequently centrifuged for 3 minutes, the organic phase was the decanted from the pellet into a clean small Kimax test tube and evaporated to dryness under a stream of nitrogen at 40 °C for approximately 45 minutes.

Derivatization reagents BSTFA, TMCS and pyridine (volumes determined using creatinine value as indicated in section 4.4.1, Table 4.2) were then added to the dry organic extract using a 100 µl Hamilton syringe. Samples were derivatized at 60 ºC for 60 minutes in a heating block. Each derivatization mixture was then transferred to a 1.5 ml GC-MS sample vial and 1µl was injected into the GC×GC/TOFMS.

4.4.3 Gas Chromatography Mass Spectrometry analysis

In this metabolomics investigation, a GC×GC-TOFMS was used for elucidating possible metabolite disturbances or biosignatures related to TB infection and treatment. GC×GC-TOFMS consists of a two dimensional gas chromatograph equipped with two columns of different polarity and mechanisms of separation (Kalinova *et al.,* 2006). A thermal modulator situated between the two columns ensures improved compound separation. The two dimensional GC is combined with a time of flight mass spectrometer (TOFMS) detector, responsible for the quantification and identification of compounds (Kalinova *et al.,* 2006). This technique also is extensively proven to be able to differentiate metabolite profiles of complex biological samples (Ralston-Hooper *et al.,* 2008; Schoeman, du Preez & Loots, 2012; du Olivier & Loots, 2012).

Chromatographic analyses of derivatised samples were done in the two-dimensional mode on an Agilent 7890A GCxGC (Agilent, Atlanta, GA) coupled to a time of flight mass spectrometer (TOFMS) (Leco Corporation, St. Joseph, MI, USA) equipped with a Gerstel

Multi Purpose Sampler (MPS) (Gerstel GmbH & co. KG, Eberhard-Gerstel-Platz 1, D-45473 Mülheim an der Ruhr). Samples (1 µL) were injected at split ratio of 1:12, and helium was used as the carrier gas at a constant flow rate of 1 mL min^{-1}. The injector temperature was held constant at 280 °C for the entire run. A Restek Rxi-5Sil MS capillary column (30m, 0.25 mm i.d., 0.25 µm d.f.) served as the primary column and compound separation was achieved by programming the primary oven at an initial temperature of 55 °C for 1 min, followed by an increase of 5 °C min to a temperature of 125 °C, then an increase of 7 °C min to a temperature of 285 °C, at which it was maintained for 4 minutes and then finally increased at a rate of 20 °C min' to a final temperature of 305 °C, at which it was maintained for a further 1 min. A Restek Rxi-17 (1 m, 100 µm i.d., 0.1 µm d.f.) column served for the second dimensional separation of the compounds. The secondary oven was programmed with an identical temperature gradient to that of the primary column, only with an offset of + 5 °C. Cryomodulation and a hot pulse of nitrogen gas of 0.7 seconds, every 3 seconds, was used to control the effluent emerging from the primary column onto the secondary column. No mass spectra were recorded for the first 480 seconds of each run, as this period was considered a solvent delay. The transfer line was held at a constant 280°C and the ion source temperature at 200°C, for the entire run.

The detector voltage was 1600 V and the filament bias -70 eV. Mass spectra were collected at an acquisition rate of 200 spectra's per second from 50-600 *m*/*z*.

4.5 Data processing

4.5.1 Deconvolution and peak identification

ChromaTOF software (version 4.32) from the Leco Corporation was used for peak finding, mass spectral deconvolution (spectral determination) and peak identification. With peak finding, the positions of all peaks in a sample including multiple components in complex co-elutions are located.

Mass spectral deconvolution is based on the fact that the mass spectrum of each analyte is uniform across the entire chromatographic peak. Deconvolution is then used to resolve the mixed mass spectra of co-eluting compounds, into accurate individual mass spectra for each analyte. Additionally, deconvolution also assigns the correct signal intensity to each analyte,

which allows for rapid, reliable quantification of co-eluting compounds. Mass spectra were deconvoluted at a Signal to Noise (S/N) ratio of 300, with a minimum of 3 apexing peaks. Once the peaks are located and the mass spectrum determined for each individual sample, a comparison is made based on mass spectrum similarities and peak positions (retention times). Using the mass fragmentation patterns generated by the MS, together with their respective GC retention times, the identities of these peaks were determined using libraries generated from previously injected standards.

4.5.2 Peak alignment

Statistical compare, an optional function on the ChromaTOF software, was used for peak alignment prior to statistical data analyses. Statistical Compare has three main features, namely: 1) peak alignment, 2) basic statistics, and 3) Fisher ratio calculations. The peak alignment feature, functions by using retention times and mass spectra to align peaks between a variety of samples. The effects of retention time shifts are also eliminated in order to compile comparative data matrixes for statistical data analyses.

The output from Statistical Compare included the following information: name of compound, retention time and peak area. A data matrix was subsequently created which included the above mentioned information, for all aligned compounds detected, for all patient urine samples analysed. Various pre-processing steps were then applied to the data.

4.5.3 Data Clean-up

A total of 782 compound peaks were detected in all the analysed patient urine samples, 315 of which could be identified by mass spectra and retention index comparisons. Prior to data normalisation, a 50% filter was applied to the peak areas in order to remove compounds which were not detected in at least 50% of samples, and removal of those compounds that showed no variation between the groups. The peak areas were subsequently normalised relative to the internal standard, to eliminate any variation which may occur due to irregularities during the organic acid extraction or injection and analysis of the samples on the GCxGC-TOFMS. The area of each peak was used to determine the relative concentration for each respective compound, using the following formula:

Relative concentration (mg/g creatinine) = area of peak/area of IS x 262.5

(262.5 is a constant, factoring in the molecular mass of creatinine)

A quality control co-efficient of variation (QC-CV) filter was then applied to the data. The coefficient of variation (CV) for all compounds detected in the QC samples analysed were calculated. The CV-value, is defined as the ratio of standard deviation to the mean of the concentrations ($CV = s/x*100$), where x is the mean and s is the standard deviation of the compounds present in the samples. There is a direct correlation between the concentration of the compound and its CV value. Compounds with large concentrations have inherently lower CV values, whereas compounds with small concentrations, have larger CV values due to larger variation (Rocke & Lorenzato, 1995). For the purpose of target compound analyses, the Food and Drug Administration (FDA) allows a CV value of 30% for GC-MS analyses. Dunn *et al*. (2011) however, describes a 20% cut-off for the GC-MS analyses of compounds more closely related to those extracted and analysed in our investigation. However, as metabolomics can be defined as the unbiased identification and quantification of all metabolites (Dunn *et al*., 2011), including being unbiased to compounds in lower concentrations, which may suffer from slightly higher CV values, a CV value of 50% was chosen as benchmark to include all compounds in the final data matrix, which was used in the statistical data analyses, for metabolite marker identification (Schoeman *et al*., 2012; du Preez & Loots, 2012). However, when evaluating the repeatability of the methodology used in Aim 1, the standard cut-off of 20%, as described by Duran et al. (2011), was used for evaluation of the data generated by this methodology.

The calculated CV values of all QC samples were used and all compounds in these QC samples, with a calculated CV value of above 50%, were removed from the entire data set, including the analysed patient samples, as these compounds represent variation/noise in the data. Following this, a zero value replacement was done, which is standard practice for subsequent multivariate statistical analyses (PCA and PLS-DA). This is accomplished by replacing zeros in the data set, with a constant calculated as half of the minimum value in the data set. This is done, as most missing values are considered to be due to a low abundance of these metabolites in the sample, (hence metabolites below the detection limit of the apparatus), and not necessarily because they are absent.

A “QC correction” step was then applied to the data set. In this step, QC samples of the various batches are used to correct for any nonbiological variation which may exist when comparing the sample batches analysed, and hence corrects for any batch effects. These differences may be attributed to small analytical variation which may occur over the duration of the analysis. This data transformation method also known as quantile equating, and corrects for any linear and nonlinear differences in distribution which may be present among batches of semiquantitative data, obtained from the same analytical method (Draisma *et al*., 2010).

Finally, samples which have been identified as outliers were removed using the following standard procedure: the sum of the concentrations of all the compounds detected in each patient’s sample was determined, the average (AVG) and standard deviation (STDEV) were calculated for all of the sum totals of each sample, and a maximum and minimum range for these were calculated using the following formulas respectively, max limit = AVG + (STDEV x 1.96) and min limit = AVG - (STDEV x 1.96). If the value of the sum total of all compounds in a sample was found to be higher or lower than the maximum or minimum limit calculated, this was then considered as an outlier, and the entire sample removed from the database. The remaining data were then pre-treated using log transformation and mean centring to compile a final dataset which was subjected to univariate and multivariate statistical analysis. This final data base contained 231 aligned compounds, for which 126 had annotated names, and 105 identified as unknowns. The high amounts of unknowns in this data base is most probably due to the drug metabolites present in the samples collected at weeks 1, 2, 4 and 26, which as of yet, have not yet been identified and hence their spectra are absent in the commercially obtained and in-house libraries used for peak identification in this study.

The datasets were now ready for advanced statistical analysis. Table 4. summarises the total aligned compounds for each time interval, as well as the amount of compounds which could be annotated by comparison of the mass spectra and retention times obtained from libraries compiled from previously injected standards. The total unidentified compounds for each time interval are also indicated.

Table 4.3: Summary of the total aligned compounds for each time interval, as well as the amount of compounds which could be annotated, or otherwise remained unidentified.

Time Point	Aligned Compounds	Annotated	Unidentified/ Unknown
Week 0	172	107	65
Week 1	167	106	61
Week 2	169	107	62
Week 4	171	107	64
Week 26	169	107	62

4.6 Statistical data analysis

Statistical data analysis was performed using the statistical package, Statistica, version 10 (Principal component analysis (PCA) and Partial-least squares discriminate analysis (PLS-DA)), as well as a web-based server, MetaboAnalyst (*t*-test, analysis of variance and ANOVA-simultaneous component analysis). MetaboAnalyst is based on the statistical program "R" version 2.10.0 and provides a wide variety of options for metabolomics data processing, data normalization, multivariate statistical analysis, graphing, metabolite identification and pathway mapping (Xia *et al*., 2009).

A variety of univariate and multivariate techniques were used to statistically analyse the data in order to reach the various aims of this study.

4.6.1 Univariate analysis

Univariate analyses are used to obtain an overview or estimated raking of potentially important features, by separately examining each variable individually, without taking the effect of multiple comparisons into account (Xia *et al*., 2009).

4.6.1.1 *t*-test

The *t*-test determines whether or not the averages of two groups are different. First a *t*-value is determined and then a *p*-value can be calculated, which is then used to establish if the difference in average is statistically significant for a single compound, between the groups (Xia *et al*., 2009). The *t*-test supports both paired and unpaired analysis. The non-parametric Mann-Whitney test were used to compare the two outcome groups at time of diagnosis, before treatment commenced as stated in Objective 2. The paired *t*-test on the other hand was applied to the data comparisons for Objective 3, comparing the two outcome groups using the calculated differences in metabolite concentrations at week 4 of treatment and time of diagnosis, due to the fact that the same patient urine was collected at these two time points, the data is therefore dependant. In both instances, a *p-value* < 0.05 was considered significant (Ellis & Steyn, 2003) and used for marker selection.

4.6.1.2 Effect size

Effect size (ES) is a simple way of quantifying the difference between two groups, by emphasising the size of the effect, rather than confounding this with sample size, and therefore it is a true measurement of the significance of the difference. Thus, it is a sophisticated, particularly valuable tool for quantifying, reporting and interpreting the effectiveness of a particular intervention, relative to some comparison, by determining how well it works in a range of contexts. ES is thus considered a measure of practical significance and is used as a univariate analytical approach to determine the importance of a variable. The ES of each compound was subsequently determined by calculating its *d*-value, using the standardised mean difference between the cured and failed groups, or rather $d = |X_1 - X_2|/S_{max}$, where X_1 and X_2 signifies the group means and S_{max}, the maximum standard deviation of two groups analysed. For parametric data, an effect size of $d < 0.2$ can be considered as being of small practical importance, whereas an effect size of $d > 0.5$ and $d > 0.8$ signifies medium and large practical significance respectively. In the case of non-parametric data, an effect size of $d > 0.5$ and $d > 0.3$ signify respectively a large and medium practical significance, whereas an effect size < 0.1 is considered as irrelevant (Ellis & Steyn, 2003).

For Objective 2, the two outcome groups were compared at time of diagnosis, and the data used for this comparison was of a non-parametric nature, hence an effect size of d > 0.3 was used to select/identify compounds/markers of importance. For Objective 3, the two outcome groups were compared on the basis of their calculated differences in metabolite concentrations at week 4 of treatment and time of diagnosis. The data used for this comparison were of a parametric nature and an effect size of > 0.5 was used to select/identify compounds/markers of importance.

4.6.1.3 Analysis of variance

Analysis of variance (ANOVA) is a parametric technique used to determine the statistical significance of the differences between the means of two or more populations or groups of values. ANOVA is therefore used as an extension of the *t*-test (de Haan *et al*., 2007). A variety of ANOVA techniques are used for statistical analysis, however, for this study the two-way ANOVA was chosen for analysis.

Two-way ANOVA is used to compare the means of populations or groups that are classified in two different ways, that is, the mean response in an experiment with two factors. Two-way within-subject and between-subject ANOVA are available for specific analysis of sample groups (Xia *et al.,* 2012). In this study, the two-way within-subject (same subjects measured over time) ANOVA was used for comparing the treatment success and failed outcome groups, over the entire 6 month treatment regimen, as described in Objective 4, with the two factors being treatment outcome and time. In the first step, ANOVA determines whether or not an interaction exists between the two factors for each variable. If an interaction is present, the differences in the mean value of the dependent variable will be identified as being statistically significant at 5% significance level ($p \leq 0.05$), which is an indication that the specific variable is influenced by both factors. In the case of an interaction between the factors, the effects of the factors in isolation are also examined to determine if a difference exist as time passes or as treatment outcome changes. Dependent variables for which group differences are statistically significant ($p \leq 0.05$) for treatment outcome but not for time are then identified. Variables which are statistically significant ($p \leq 0.05$) for the time factor but not for the treatment outcome factor, are subsequently also identified. Thus the two-way

within-subject ANOVA used, identified three sets of dependent variables for which group means differ significantly, however, the variables identified to be statistically significant for time, were not used in this study.

4.6.2 Multivariate analysis

Univariate techniques are easily understandable; however, their applications are often limited when dealing with large data sets containing large amounts of highly correlated variables. For this reason, multivariate methods are also used and highly depended on for marker selection during metabolomics analyses. As the name implies, multivariate analysis consists of a collection of statistical techniques, that can be used when various measurements are made on each variable, in one or more samples (Rencher, 2002).

4.6.2.1 Principal component analysis

Principal component analysis (PCA), is a multivariate, unsupervised classification method, which aims to determine whether or not a natural differentiation between experimental groups exists. PCA is based on a mathematical procedure that transforms possibly related variables (metabolites) into a smaller number of unrelated variables, known as the principal components (PCs). PC 1 accounts for the most variance in the data and each following PC (PC 2, PC 3, etc.) accounts for the next highest variance of the remaining data. Using one PC per axis, a PCA is visualised as a scores plot, and the metabolites best describing the variation in the data, are ranked according to their respective PCA modelling powers. A modelling power > 0.5, is stated by Brereton *et al.,* (2003), to signify an important variable in the PCA projection and subsequently used a cut-off for metabolite marker selection, for those compounds best explaining the variation between the groups compared. PCA was subsequently used in the current study to determine whether or not a natural differentiation existed between patients with a failed and cured treatment outcome; 1) at time of diagnosis (Week 0), as described in Objective 2, and 2) to the data set comparing the two outcome groups using the calculated differences in metabolite concentrations at week 4 of treatment and time of diagnosis (Objective 3), in order to identify metabolite markers explaining the underlying differences/variation between the compared groups, at these time points.

4.6.2.2 Partial-least squares discriminate analysis

Partial-least squares discriminate analysis (PLS-DA) is a multivariate, supervised classification method, which uses various linear regression techniques in order to find the direction of maximum covariance between a data set and a class membership (Xia *et al*., 2009). PLS-DA, uses this mathematical approach, to summarise the original variables, into fewer variables, also known as scores, using their weighted averages. The PLS-DA thus provides a second set of important metabolites i.e. variables important in projection (VIPs), defined as the weighted sum of squares of the PLS loadings, which takes into account the amount of explained class variance of each component. Thus VIPs are the components or metabolites which differ the most between the groups and best explain the inter-group variation.

Similarly to the PCA analysis, the same two data sets (1) week 0 (at time of diagnosis) and (2) the data set comparing the two outcome groups, using the calculated differences in metabolite concentrations at week 4 of treatment and time of diagnosis, were also subjected to PLS-DA analysis. VIPs were calculated and used to select those variables/metabolites of importance. Variables with a VIP > 1, were identified and selected as potential markers (Chong & Jun, 2005).

4.6.2.3 ANOVA-simultaneous component analysis

ANOVA-simultaneous component analysis (ASCA), can be viewed as a multivariate extension of the univariate ANOVA approach. In ASCA, the Multivariate ANOVA (MANOVA) approach is combined with PCA. ASCA can therefore examine the variation induced by known influential factors for a number of metabolites simultaneously. ASCA is designed to identify dominant patterns of variation associated with each factor and can also be used to investigate factor interactions. Firstly, the overall observed variation is partitioned into the variation associated with each factor and factor interaction. PCA is then applied to each partition to uncover the internal structure that best accounts for the observed variation. A PCA score plot is subsequently compiled to visualise the variation explained by the associated components. Finally, the leverage and squared prediction errors (SPE)

associated with each variable is calculated. The leverage value is considered a measure of the contribution of the variable to the fitted ASCA model, subsequently a higher leverage indicates more importance (Xia & Wishart., 2011; Smilde *et al*., 2005). The SPE value is an evaluation of the goodness of fit of the ASCA model to a specific variable, in which case, a higher SPE value means less fit. Finally variables with a leverage value between 0 and 1 (Nueda *et al*., 2010), are considered well modelled variables and identified as being most significant and used for marker selection.

ASCA was applied to a data set, containing the analysed data, of all patients in the cured and failed outcome groups, for which a patient urine sample was collected at all the described time points (Objective 4). The two factors selected for analysis included: time and treatment outcome. Two sets of important variables with respect to treatment outcome and the interaction between time and treatment outcome were identified by means of ASCA analysis, and again variables associated to be changed over time only, were not used in this study.

4.7 Summary of experimental design

The broad outline of this experiment is shown in the figures below. Figure 4.2, indicates the experimental plan proposed to answer the research question, as formulated in the aims and objectives. The experimental design is divided into two separate sections, (1) testing the repeatability of the methodology and analytical apparatus used - Aim 1, Objective 1 (shown on the right side of this flow), and (2) the generation of metabolomics data, using patient collected urine samples - Aim 2, Objectives 2-4 (shown on the left side of this flow).

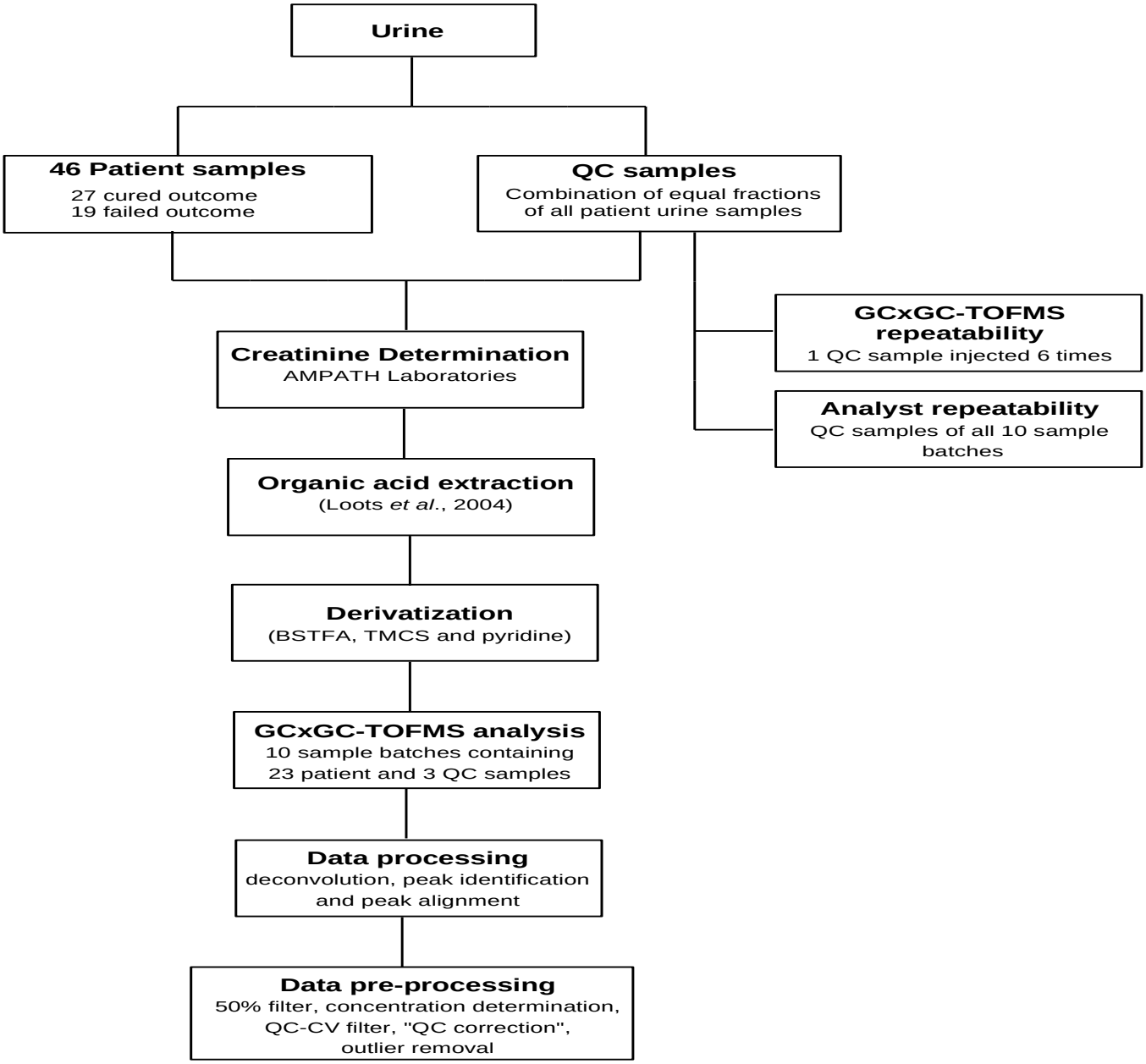


Figure 4.2: Schematic representation of the experimental design used for this metabolomics study.

Subsequently, a summary of the statistical approach in order to achieve the research Aim 2 as described above, is given in Figure 4.3

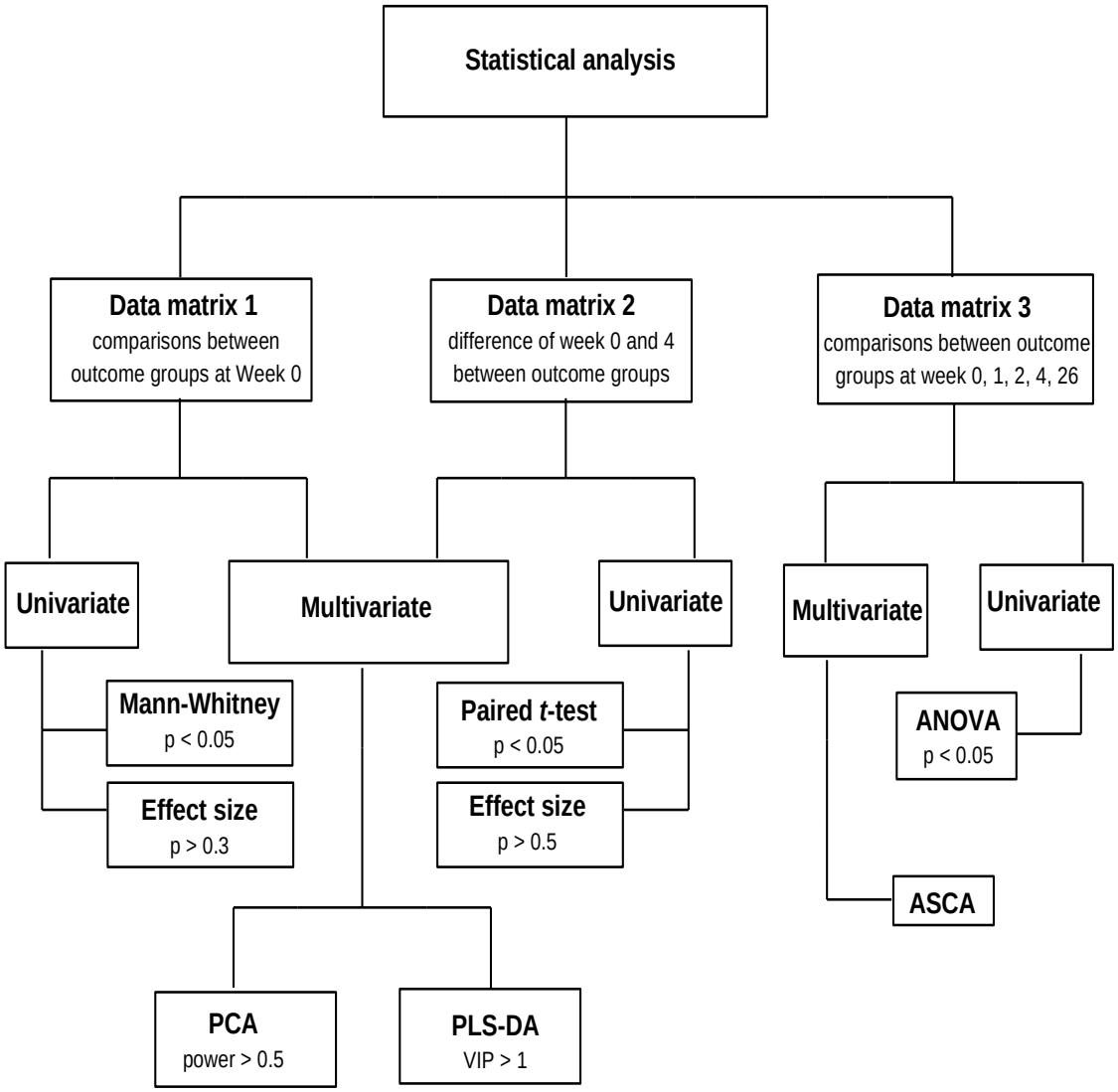


Figure 4.3: Work-flow of the statistical methods used in this metabolomics study

Chapter 5

Results and Discussion

5.1 Repeatability and Reliability of Data

In this section, the repeatability of the GCxGC-TOFMS, as well as that of the extraction procedure / analyst's ability to perform this repeatability is investigated. Additionally, possible batch effects are determined, which, as described in the methods section, will be corrected for using quantile equating (section 4.5.3).

5.1.1 Repeatability of GCxGC-TOFMS and extraction method/analyst

In order to evaluate the GCxGC-TOFMS analytical repeatability for the analysis of the prepared extracts, one QC sample was extracted using the organic acid extraction analysis and derivatization procedure as described in section 4.4.2. The QC sample was injected repeatedly (6 times) onto the GC×GC-TOFMS. Data processing as described in section 4.5, was applied and a final dataset of the 6 sample repeats, was subsequently used to calculate the CV values (section 4.5.3) for the 1) compound areas and 2) calculated relative concentrations, for all the compounds detected in the prepared extracts analysed. The repeatability of the analytical technique was then determined by evaluation of the calculated CV values with reference to the prescribed cut-offs available in literature.

The calculated CV values for ten compounds, representative of various compound classes detected at regular retention time intervals throughout the total chromatographic run, are presented in Table 5.1. The CV values for these compounds confirm that the GCxGC-TOFMS provides repeatable results for the extracts prepared, when compared to the predetermined cut-offs of 20% as described by Dunn *et al*. (2011), and hence its suitability for the analysis of these extracts for metabolomic applications.

Table 5.1: Coefficients of variation (CV) values, calculated for ten compounds using compound area and the relative concentration respectively, after GCxGC-TOFMS analysis.

Compound	Retention time (sec)	Area	CV Value % (Area)	Concentration (mg/g creatinine) ± STDEV	CV value % (Concentration)
N-Benzoyl glycine	1728	28529310	14.18	4437.132 ± 533.979	12.03
4-Hydroxy benzeneacetic acid	1530	3325421	14.35	514.698 ± 27.345	5.31
3-Hydroxyhippiric acid	1911	3062145	10.59	479.154 ± 42.309	8.83
Hydroxyacetic acid	756	1181249	13.57	182.720 ± 3.917	2.14
4-Hydroxy-3-methoxy benzoic acid	1650	792382.3	14.47	122.726 ± 8.239	6.71
Ethanedioic acid	858	726210.4	10.25	112.811 ± 6.132	5.44
2-Hydroxypropanoic acid	726	534825.8	10.88	82.996 ± 3.493	4.21
5-Oxo-proline	1404	507298.1	15.81	78.323 ± 3.008	3.84
2-Hydroxy pentanedioic acid	1464	452730.2	14.97	70.197 ± 6.625	9.44
Glucuronic acid	1830	438285.8	15.47	67.89 ± 5.6096	8.26
3-Phenylbutyric acid (IS)	1296	1698857	14.54		

IS = Internal Standard

Similarly, the repeatability of the extraction method/analyst's ability for performing these analyses in a repeatable manner was also determined. One QC sample, was repeatedly extracted and derivatized multiple times (n = 10) as described in section 4.4.2, and the analytical data collected after GC x GC/TOFMS analysis processed as described in section 4.5. The repeatability of the extraction method was then determined by evaluation of the calculated CV values with reference to the prescribed cut-offs available in literature.

Once again, the CV values were calculated for the same ten compounds described above, which are representative of the various compound classes, and detected at regular retention time intervals throughout the total chromatographic run. This information is presented in Table 5.2. The average CV values for these compounds confirm that the extraction method /analyst's ability for performing this, is sufficiently repeatable for application to this

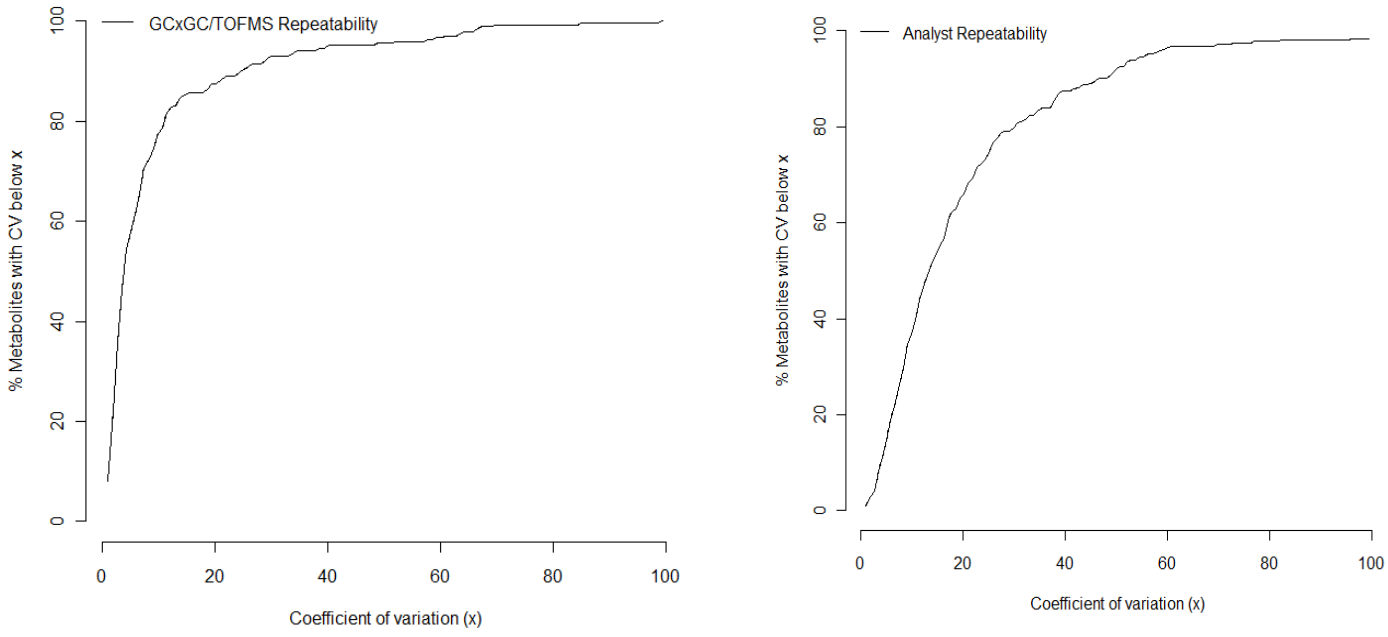
metabolomics investigation, considering the predetermined cut-offs in the literature of 20% (Duran *et al*., 2011).

Table 5.2: Coefficients of variation (CV) values, calculated for ten compounds detected in all QC samples after organic acid extraction analysis and derivatization, using compound area and relative concentration respectively.

Compound	Retention time (sec)	Area	CV value % (Area)	Concentration (mg/g creatinine) ± STDEV	CV value % (Concentration)
N-Benzoyl glycine	1730	2653735	18.37	150.082 ± 25.17	16.77
4-Hydroxy benzeneacetic acid	1531	2075822	15.62	60.007 ± 8.445	14.07
3-Hydroxyhippiric acid	1910	1721122	16,11	49.707 ± 6.857	13.80
Hydroxyacetic acid	758	1323458	15.52	38.304 ± 5.719	14.93
4-Hydroxy-3-methoxy benzoic acid	1650	413677.5	17.88	25.974 ± 3.882	14.95
Ethandioic acid	858	494537.3	19.15	14.093 ± 2.531	17.95
2-Hydroxy propanoic acid	725	413956.5	20.19	11.880 ± 2.188	18.41
5-Oxo-proline	1404	379250.7	21.66	10.909 ± 1.941	17.80
2-Hydroxy pentanedioic acid	1464	314276.6	16.66	9.088 ± 1.055	11.61
Glucuronic acid	1830	215750.9	18.50	8.960 ± 1.402	15.64
3-Phenyl butyric acid (IS)	1296	9143188	14.32		

When evaluating these results, one should keep in mind, that the CV of a compound is also directly correlated to its detected concentration, which means that a compound detected in a higher concentration, would inherently be accompanied by a comparatively lower CV value, as opposed to one occurring in a lower concentration, accompanied by a higher CV value (Rocke & Lorenzato, 1995). As previously described, metabolomics, per definition, is geared towards analysing, evaluating and including, all compounds, in an unbiased manner, including those in lower concentrations, which would most likely be accompanied by higher CV values. For this reason, a larger CV cut-off (50%) as an inclusion criteria for compounds in the final data matrix, prior to statistical analysis for metabolomics biomarker identification, has been described (Schoeman *et al*., 2012; du Preez & Loots, 2012), and will be used in the subsequent analyses, to address the research aims in this investigation.

When using these cut-offs, an average of 95.59% of all the compounds detected from the 6 QC sample repeats used for determining the GC x GC-TOFMS repeatability, fell below this 50% CV value, which somewhat astonishing, and further proof of the excellent analytical repeatability of the GCxGC-TOFMS, for the majority of the compounds extracted using the previously described extraction method (Figure 5.1). Similarly, the extraction method/analyst's ability to perform this in a repeatable fashion, also showed 93,7% of all compounds analysed to have a CV < 50% (Figure 5.1).



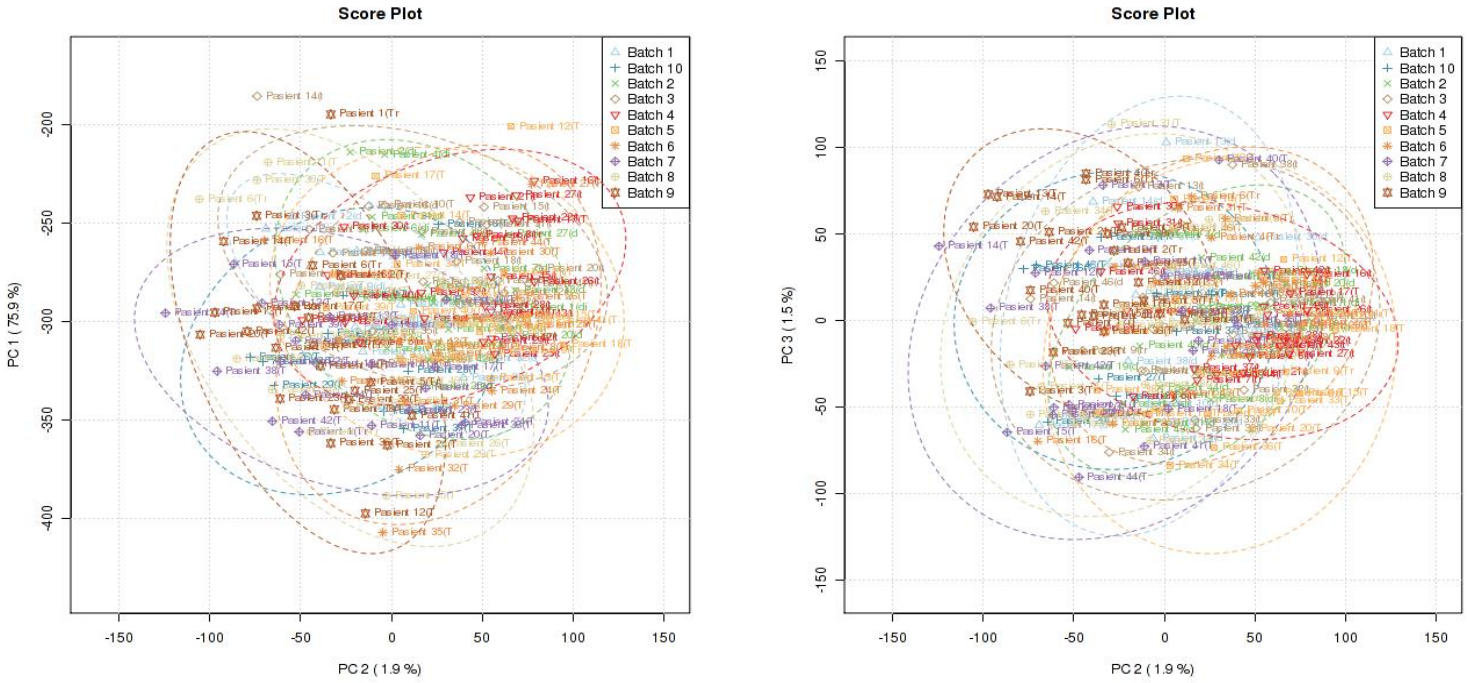
a) b)

Figure 5.1: Distribution of the coefficients of variation (CV) values determined from the relative concentrations of all the compounds detected using the QC samples as a measure for a) GCxGC/TOFMS repeatability and b) Extraction method/ Analyst repeatability, subsequent to an organic acid extraction, and GCxGC/TOFMS analysis.

5.1.2 Batch effect

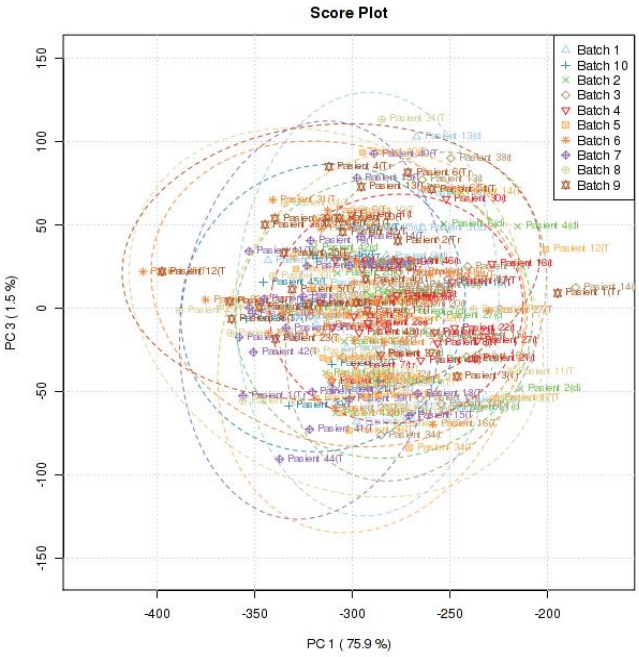
Luo *et al*, (2010) defines, the term batch effect as: the systematic non-biological differences between batches of samples in experimental groups, as a result of various conditions or confounding factors which arise during sample handling or during analytical analyses. It is also stated that batch effects could be reduced or in most cases avoided, if an experimental design is well set out and the necessary precautions are taken, however, in some circumstances, certain batch effects are unavoidable. Since these batch effects are due to non-biological circumstances, they could potentially confound true biological differences between sample groups.

In this study, as previously described, urine samples from 46 patients clinically diagnosed with TB, were collected at various time points during the 6 months DOTS treatment regimen. The 210 collected patient samples were extracted and analysed (section 4.4.2) in ten batches. Quantile equating was done as described in section 4.5.3., in order to remove any potential batch effects which may have occurred. In order to confirm that this was successful, a PCA was done using the corrected data, to visualise whether or not any differentiation of the groups exists, due to a batch effect. Figure 5.2 shows the 2D PCA score plots, visualising all samples analysed in each batch.



a)

b)



c)

Figure 5.2: 2D score plots of a) PC 1 vs. PC 2, b) PC 2 vs. PC 3 and c) PC 1 vs. PC 3 of the 10 batches in which the 210 patient urine samples were analysed. The percentage cumulative variance explained by the first two PCs (R^2X cum) was 13.4%, of which PC 1 explained 7.0% and PC 2 explained 5.4%.

It is apparent from the illustration provided by Figure 5.2 and the percentage cumulative variance explained, that there is no differentiation between the batch groups, and that all batches overlap, indicating that no batch effect exists in the analysed metabolite data, and any differentiation they may later occur when comparing the cured and failed treatment outcome groups, are due to a metabolite difference between these groups due to an underlying biological effect.

5.2 Metabolomics investigation of TB treatment outcome

We investigated three research questions, under the following sections:

5.2.1 Metabolomics comparisons of cured and failed treatment outcomes at time of diagnosis (before treatment), in order to identify potential biomarkers, possibly explaining variation in these groups due to their varying response to the TB disease state.

5.2.2 Metabolomics comparisons of cured and failed treatment outcomes during treatment, in order to identify potential biomarkers possibly explaining variation in these groups due to their varying response to the treatment regimen.

5.2.3 Metabolomics comparisons of cured and failed treatment outcomes over entire study duration, considering all time points simultaneously, over the 6 month treatment regimen, using a more conventional statistical approach, in order to identify new biomarkers or confirm the previously identified biomarkers, better characterising the differences which may exist between those who were eventually cured and those where treatment failed, and the causes for these differences.

We anticipate that the metabolomics information generated from this investigation, will lead to a better understanding of as to why some individuals respond to treatment, and why others don't.

5.2.1 Metabolomics comparisons of cured and failed treatment outcomes at time of diagnosis

Prior to univariate and multivariate statistical analysis, a variety of prepossessing steps were applied to the collected GCxGC-TOFMS data, as described in section 4.6.3. For the application of multivariate statistical analysis, the dataset was cube root transformed and for univariate statistical analysis, a log transformation function was applied. PCA, PLSDA, Mann-Whitney test and effect size calculations were subsequently applied to the processed dataset. This dataset contained 172 aligned compounds, 107 of which could be annotated by comparison of the mass spectra and retention times of each compound, to that of libraries compiled from previously injected standards.

5.2.1.1 PCA differentiation between successful and unsuccessful treatment outcome groups at time of diagnosis

A PCA was performed on the GCxGC/TOFMS generated dataset in order to determine whether or not a natural separation exists between the two treatment outcome groups at time 0, before any treatment commenced. Using the 172 variables in the processed data set described above, a clear differentiation of the two groups was achieved by the first two PCs of the PCA (Figure 5.4). The total amount of variance explained by the first two PCs (R^2X cum) was 93.4%, of which PC 1 explained 91.8% and PC 2 explained 1.6%. This clear differentiation achieved, can be ascribed to the differences in the extracted metabolite profiles of the individual patients in each group. This result, is not only astounding from a biological point of view, considering that we may potentially explain why certain individuals would respond to treatment, and others not, due to differences in their metabolome, prior to receiving treatment (most probably due to variation in their genome), but also from a diagnostic perspective, potentially being able to determine if the patient will respond to the treatment or not, even before treatment commences.

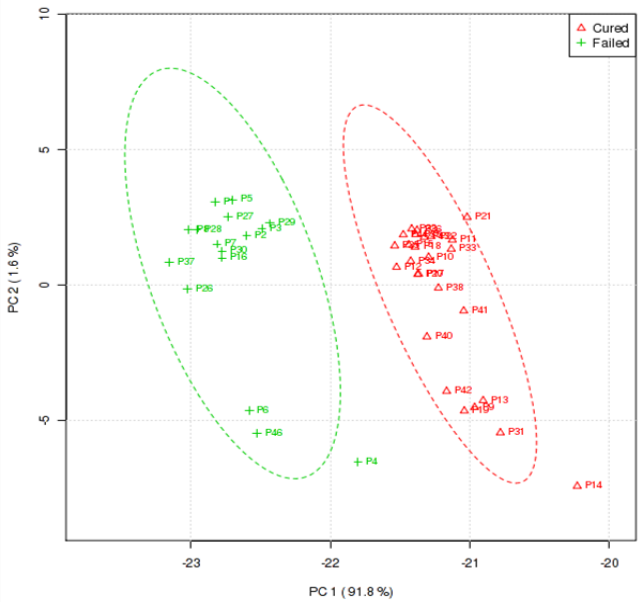


Figure 5.3: PCA scores plot illustrating PC1 vs. PC2 of the successful and failed treatment outcome group at time of diagnosis, subsequently to organic acid extraction and GCxGC/TOFMS analysis, indicating a differentiation of the two outcome groups. The variance explained by each PC is given in parenthesis.

5.2.1.2 Metabolite marker identification

In a metabolomics research approach, the primary aim is not only to determine whether or not differentiation of various sample groups exists, which is based on varying metabolite profiles of the compared groups, but also the identification of those metabolite markers, which best described the variation seen between the sample groups. These metabolite markers are subsequently used to better characterise the investigated sample groups and also provide important information which is needed to answer a biological question.

In our quest to identify metabolite markers, results from univariate and multivariate statistical methods were combined in order to compile a list of metabolite markers which show most significant variation between the two treatment outcome groups. As illustrated in Figure 5.3, a natural differentiation between the two groups exists and all the metabolites associated or responsible for this differentiation were subsequently identified. All metabolites with a PCA modelling power > 0.5 were selected (Brereton, 2003), representing those

metabolites/variables best describing the variation seen in the unsupervised, multivariate PCA projection. PLS-DA on the other hand, is a supervised multivariate method, and was used to determine the VIP values of all the metabolites present in the two sample groups. Metabolites with a VIP value > 1 were selected as potential markers (Chong & Jun, 2005) and represent those metabolites which vary the most between the compared sample groups. The PLS-DA model used, had a modelling parameter, R_2Y (cum) of 93.5%, indicative of the total explained variation of the response, Y. The cross-validated variation, Q_2 (cum) explained by the response Y, was 48.1%.

A number of univariate statistical methods were also used from further metabolite selection, and included effect size and Mann-Whitney test calculations. Metabolites with an effect size > 0.3, and a Mann-Whitney test p-value < 0.05 were considered as being most significant. A detailed explanation for these statistical methods are given in Chapter 3, section 4.7.

Illustrated in Figure 5.5, using the above statistical methods, the list of metabolite markers were compiled from a combination of the selected PCA and PLS-DA markers, having an ES > 0.3. The *p*-values for each of the selected metabolite markers are also reported (Table 5.3), however, not used for metabolite marker selection, as this statistical method is highly dependent on same size.

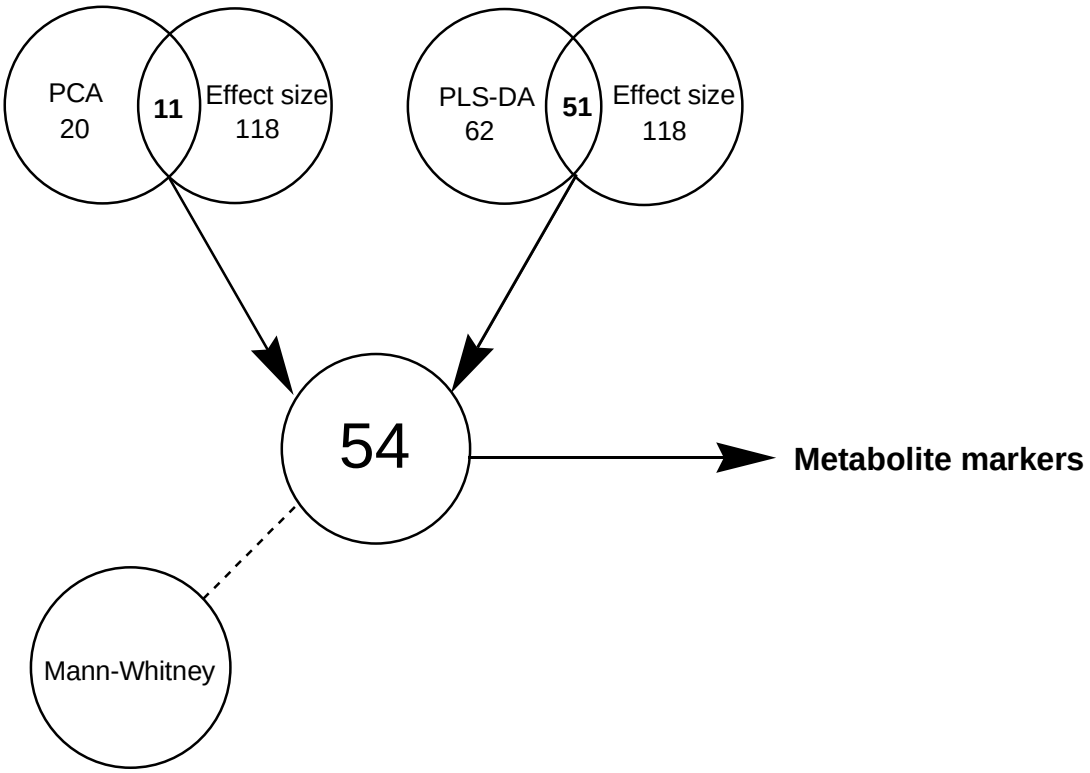


Figure 5.4: Statistical approach used to identify metabolite markers which best explain the variation between the successful and failed outcome at time of diagnosis.

Of the 54 compounds selected as metabolite markers using this above selection approach, 39 were annotated using libraries prepared from injected standards, by comparison of their mass spectra and retention times. The PCA power, PLS-DA VIP, average concentration, standard deviation, effect size and Mann-Whiney test (*p*-value), for each of the 39 metabolites, are provided in Table 5.3. These metabolites will subsequently be used in order to explain the biological variation between those individuals who were successfully cured, using the standard TB treatment protocol, vs those who weren't, at time of diagnosis, before any treatment was administered.

Table 5.3: Metabolite markers best describing the variation between the successful and faild treatment oucome groups

Metabolite Name			Cured treatment outcome	Failed Treatment outcome		
	PCA Power	PLSDA VIP	Average concentration ± standard deviation (mg/g creatinine)	Average concentration ± standard deviation (mg/g creatinine)	Effect Sizes	Mann-Whitney (p-value)
Markers of *M. tuberculosis* metabolism						
p-Hydroxybenzoic acid	0.533	2.395	4.788 ± 3.426	9.966 ± 7.014	0.738	0.001
Ribitol	0.184	1.166	1.110 ± 1.167	2.454 ± 2.091	0.643	0.184
Markers associated with an altered host metabolism						
3-Hydroxy-dodecanedioic acid	0.678	1.313	5.102 ± 8.750	2.313 ± 4.299	0.319	0.319
Homovanillic acid	0.609	1.238	17.142 ± 4.644	22.268 ± 7.318	0.700	0.007
o-Hydroxyphenylacetic acid	0.582	1.709	3.015 ± 1.989	5.293 ± 2.408	0.946	0.011
3-Methyladipic acid	0.429	1.948	7.397 ± 3.871	12.63 ± 6.546	0.799	0.007
Sebacic acid	0.417	1.446	1.243 ± 1.443	2.560 ± 2.594	0.508	0.131
2-Deoxypentonic acid	0.415	1.246	3.977 ± 2.480	5.697 ± 3.158	0.545	0.118
Adipic acid	0.386	1.068	0.726 ± 1.01	1.369 ± 1.770	0.363	0.173
Suberic acid	0.377	1.363	3.234 ± 2.933	4.576 ± 1.573	0.458	0.022
Suberic acid	0.377	1.363	3.234 ± 2.933	4.576 ± 1.573	0.458	0.022
Glucuronic acid	0.330	1.188	5.888 ± 2.335	8.307 ± 3.027	0.799	0.013
3,4-Dihydroxybutyric acid	0.289	1.114	9.095 ± 4.353	11.729 ± 3.557	0.605	0.029
Quinolinic acid	0.287	2.537	5.312 ± 4.523	9.447 ± 5.609	0.737	0.003
Cis,cis-4,7-Decadiene-1,10-dioic acid	0.275	1.204	9.188 ± 6.432	12.113 ± 5.433	0.455	0.221
Pimilic acid	0.233	1.602	2.817 ± 2.263	4.967 ± 2.733	0.786	0.007
2-Methyl-3-hydroxybutyric acid	0.213	1.453	3.703 ± 2.560	5.486 ± 2.829	0.630	0.083
3-Hydroxysebacic acid	0.197	1.649	1.487 ± 2.040	3.716 ± 4.556	0.489	0.197
3-Hydroxyvaleric acid	0.169	1.011	2.090 ± 2.263	3.799 ± 5.204	0.328	0.169
Parabanic acid	0.152	1.543	2.254 ± 2.960	3.584 ± 3.456	0.385	0.057
2-Deoxyribonic acid	0.137	1.441	0.457 ± 0.585	1.079 ± 0.976	0.637	0.061
Azelaic acid	0.133	1.541	4.699 ± 4.832	6.464 ± 5.084	0.347	0.078

4-Hydroxyphenyllactic acid	0.129	1.065	4.557 ± 5.684	6.723 ± 6.879	0.315	0.129
Pipecolic acid	0.096	1.173	0.892 ± 1.216	0.427 ± 0.414	0.383	0.096
4-Hydroxybutyric acid	0.064	1.207	0.316 ± 0.506	0.544 ± 0.631	0.362	0.011
Methylsuccinic acid	0.051	1.590	3.833 ± 2.337	6.580 ± 4.079	0.673	0.003
Tiglylglycine	0.048	1.343	1.043 ± 1.539	1.758 ± 1.798	0.398	0.048
Markers of exogenous origin						
Vanillic acid	0.512	1.644	7.697 ± 4.636	11.249 ± 4.919	0.722	0.047
Furan-2,5-dicarboxylic acid	0.464	1.069	4.836 ± 8.193	8.051 ± 10.467	0.307	0.307
2-Furoylglycine	0.392	1.844	6.540 ± 5.889	10.421 ± 8.96	0.433	0.132
5-Hydroxymethylfuroic acid	0.360	1.695	10.823 ± 12.043	16.049 ± 15.591	0.335	0.203
Benzoic acid	0.346	1.337	0.735 ± 0.897	1.366 ± 1.001	0.631	0.001
Phenylacetic acid	0.288	1.567	0.316 ± 0.826	0.903 ± 1.089	0.539	0.008
3-Hydroxyhippuric acid	0.233	1.647	18.201 ± 8.473	23.661 ± 10.097	0.541	0.057
Dihydroferulic acid	0.123	2.603	3.007 ± 3.004	8.365 ± 9.722	0.551	0.067
4-Methylcatechol	0.123	1.255	7.147 ± 5.968	10.299 ± 7.413	0.425	0.123
3-(3-hydroxy-phenyl)-3-hydroxypropionic acid	0.049	2.337	20.766 ± 15.566	28.086 ± 16.371	0.447	0.027
Oxalic acid	0.047	1.564	13.704 ± 6.299	19.932 ± 8.463	0.736	0.03
Unclassified						
N-Formyl-glycine	0.287	1.447	0.826 ± 1.307	1.566 ± 1.483	0.499	0.034
3-Hydroxy-2,3-didehydrosebacic acid	0.017	2.682	2.478 ± 2.815	6.450 ± 6.240	0.637	0.013

These metabolite markers were broadly classified and discussed according to their various origins, including 1) those associated with the *M. tuberculosis* infection, 2) an altered host response due to the infection and 3) metabolites of exogenous origin, and will subsequently be discussed under these headings. These metabolites and their origins are summarised in Figure 5.6.

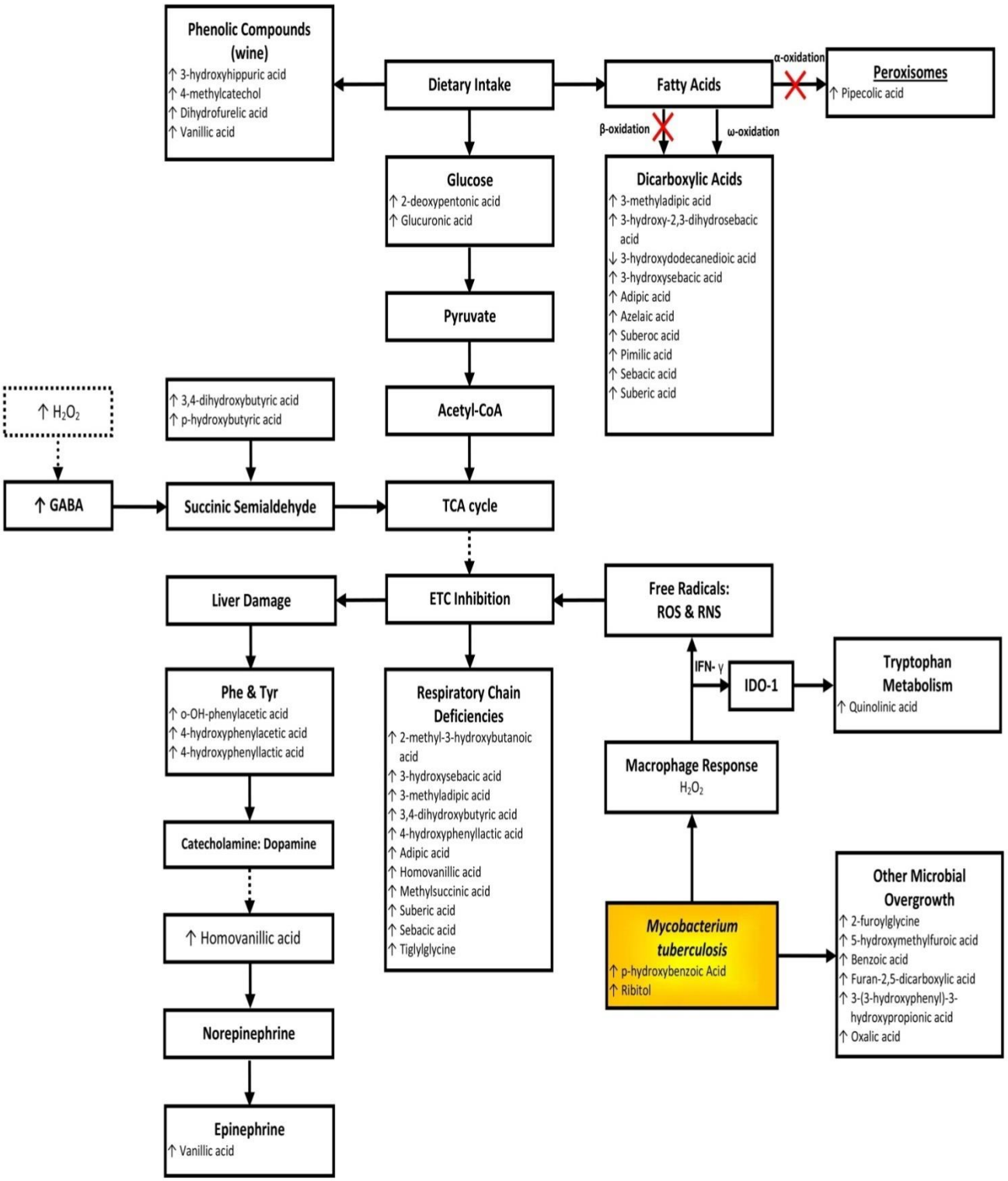


Figure 5.5: Schematic representation of elevated concentrations of the various metabolites found in the urine of the treatment failure group. This diagram indicates the metabolites associated with the *M. tuberculosis* infection, an altered host metabolism due to the infection and metabolites of exogenous origin.

5.2.1.2.1 Markers associated with *M. tuberculosis* metabolism

Among the markers identified, the two compounds, ribitol and *p*-hydroxybenzoic acid, shown to be elevated in the treatment failure group, are of *M. tuberculosis* origin directly. Gram-positive microorganisms have thick layers of peptidoglycan, which provide support to withstand the tugor pressure exerted on their plasma membranes. In these layers of peptidoglycan, threads of long anionic polymers, known as teichoic acids, are also located, which are composed of glycerol phosphate, glucosyl phosphate or ribitol phosphate repeats (Silhavy *et al*., 2010). Due to the fact that *M. tuberculosis* is also classified as a Gram-positive microorganism (Willey *et al.,* 2008), the assumption can be made that the ribitol found in the urine of these TB positive patients, most likely originates from the ribitol phosphate repeats in the cell wall of these bacteria. *p*-Hydroxybenzoic acid on the other hand, is synthesised in *M. tuberculosis* by means of the chorismate pathway known to be present in these bacteria. Chorismate plays a crucial role in the biosynthesis of a range of compounds that are essential for both the physiology and pathogenicity of these bacteria. Chorismate, is converted by chorismate pyruvate-lyase to eventually produce *p*-hydroxybenzoic acid (Stadthagen *et.al.*, 2005). The increased concentrations of these metabolites in the treatment failure group, could possibly be attributed to a greater bacterial load, however, in the patient information provided, the results of the sputum smear microscopy showed no indication of a greater bacterial load to be present in the treatment failed group (Appendix1, Table1), however sputum smear microscopy is not accurate to determine bacterial load.

5.2.1.2.2 Markers associated with an altered host response to TB

It is already well known, that infection with *M. tuberculosis,* results in a host macrophage response, engulfing the invading bacteria. Once activated by interferon gamma (IFN-γ), macrophages generate antimicrobial radicals, including reactive oxygen species (ROS) and reactive nitrogen species (RNS) (Hu & Coates, 2009). The production of ROS is due to macrophage NADPH oxidase interaction with molecular oxygen, to form superoxide, which is finally converted to toxic hydrogen peroxide (H_2O_2) and hydroxyl radicals. RNS in turn is produced by the inducible nitric oxide synthase and a highly toxic product; peroxynitrite, which may also be generated from the reaction between nitric oxide and superoxide (Hu & Coates, 2009). This is considered the primary mechanism by which the host attempts to

eliminate the invading *M. tuberculosis.* It is therefore crucial to the host to be able to produce and respond to IFN-γ in order to maintain control over the infection. Indolamine 2,3-dioxygenase-1 (IDO-1), is considered a well-defined target gene of IFN-γ signaling and is also the rate-limiting enzyme of tryptophan metabolism, which is facilitated along the kynurenine pathway to produce quinolinic acid and picolinic acid (Smith *et al*., 2001). It has been found that *M. tuberculosis* infection in the host, promotes upregulation of IDO-1 expression, which leads to an increase in tryptophan catabolism and as a result an increase of quinolinic acid rather than picolinic acid is produced in the infected macrophages (Blumenthal *et al*., 2012). Quinolinic acid was found to be elevated in the urine of the treatment failed group, which may be associated with a greater macrophage response in the failed group, however, not induced due to a higher bacterial load, as indicated by the sputum smear results (not accurate to determine bacterial load), and possibly due to some other host or bacterial phenomenon.

Furthermore, ROS is also well known to induce inhibition of the electron transport chain (Loots *et al.,* 2004), accompanied by a decline of mitochondrial function (Fosslien, 2001) as a result of a defective flow of electrons due to impaired functionality of electron transfer flavoprotein dependent enzymes (Reinecke *et al*., 2012). A marked reduction in mitochondrial production of ATP consequently results. In a study conducted by Reinecke *et al*. (2012), a metabolomics approach was used to investigate the urinary organic acid profiles of children suffering from respiratory chain deficiencies (RCDs). A variety of metabolites were identified which could serve as a characteristic biosignature for RCDs. Eleven of the metabolite markers identified in the urine of the treatment failure patient group in our metabolomics study are identical to those metabolites identified in patients suffering from RCDs, and include: methylsuccinc acid, adipic acid, suberic acid, sebacic acid, 3-methyladipic acid, 3-hydroxysebacic acid, tiglylglycine, 2-methyl-3-hydroxybutyric acid, 4-hydroxyphenyllactic acid, homovanillic acid and 3,4-dihyroxybutyric acid. This strongly suggests that the patients of the treatment failure group are suffering from a mitochondrial/ respiratory chain deficiency of some sort, resulting in the lack of response to the TB treatment.

As explained by Reinecke *et al*. (2012), the reduced mitochondrial production of ATP can be directly compensated for by the upregulation of carbohydrate, protein and triacylglycerol catabolism, by the action of various neurotransmitters such as dopamine and/or hormones

like vasopressin, glucagon and adrenaline (Reinecke *et al*., 2012). The homovanillic acid and vanillic acid detected in elevated concentrations in the treatment failure group, is indicative of elevated levels of the neurotransmitters dopamine (Frankenhaeuser *et al*., 1986) and adrenaline respectively. (Ebinger & Verheyden, 1976). Glucuronic acid and 2-deoxypentonic acid, were also detected in elevated concentrations in the urine of the treatment failure group, and associated with increased blood glucose levels (Lawson *et al*., 1976), which is most likely fuelled by the above mentioned increased adrenalin, subsequently lowering insulin. It has also been found that insulin resistance is associated with severe infection of *M. tuberculosis,* reportedly also leading to the elevated blood glucose levels in these TB patients (Faurholt-Jepsen *et al*., 2011), and further associated with the high diabetes prevalence among TB (Balakrishnan *et al*., 2012). It is well known, that severe oxidative stress, results in pancreatic damage, subsequently leading to a diabetic disease state (Loots *et al*., 2011).

Additionally, the elevated 3,4-dihydroxybutyric acid and 4-hydroxybutyric acid observed in the urine of the treatment failed group, are associated with an increase in γ-amminobytyric acid (GABA) metabolism. The elevated levels of H_2O_2 in the treatment failure group due to the proposed respiratory chain deficiency, cause impaired Cl^- gradients and a reduced efficiency of the $GABA_A$ receptor to take up the neurotransmitter, GABA (Sah & Schwartz-Bloom, 1999), and subsequently succinic semialdehyde. Additionally, inhibition of the ETC and Krebs cycle, would most likely also result in the accumulations of succinic semialdehyde, consequently resulting in the synthesis of 4-hydroxybutyric acid by means of the enzyme succinic semialdehyde reductase, and eventually 3,4-dihydroxybutyric acid (Pearl *et al*., 2011), further confirming our hypothesis of an impaired respiratory chain in the treatment failure group.

Impaired mitochondrial β-oxidation, is also seen to occur in patients with a dysfunctional ETC or RCD (Watmough *et al*., 1989; Reinecke *et al*., 2012). This is thought to be due to a decrease in electron transfer flavoprotein (ETF), as well as an increase electron transfer flavoprotein: ubiquinone oxidoreductase (ETF-QC) activity, due to a deficient flow of electrons past Coenzyme Q_{10} (Watmough *et al*., 1989, Reinecke *et al*., 2012). It is also possible that an impaired β-oxidation, results from a decrease in the NAD^+ levels (which are needed by β-oxidation) due to a dysfunctional ETC (Pessayre *et al*., 1999). An increase formation of the dicarboxylic acids, azelaic acid, pimelic acid, adipic acid, suberic acid,

sebacic acid, 3-methyladipic acid, cis,cis-4,7-Decadiene-1,10-dioic acid and 3-hydroxy-dodecanedioic acid were observed in the urine of both the RCD patients described by Reinecke *et al*. (2012) and the treatment failure group in our investigation. These are formed by ω-oxidation from their respective short- and medium-chain fatty acid-CoA precursors, comparable with the observation of impaired mitochondrial fatty acid oxidation as reviewed by Duran. (2005). Furthermore, the elevated levels of the ω-oxidation product 3-hydroxyadipic acid, again illustrates the involvement of NAD redox state. Elevated butanoyl-CoA, as a result of reduced short chain acyl-CoA dehydrogenase activity, may also produce ethylhydracrylic acid and methylsuccinic acid (Sweetman & Williams, 2001), explaining the elevated concentrations of the latter in the treatment failure group. Additionally, tiglylglycine, 2-methyl-3-hydroxybutyric acid and 3-hydroxyvaleric acid, were also elevated in the urine of the treatment failure group and are all associated with abnormalities of the isoleucine metabolism, of which, tiglylglycine and 2-methyl-3-hydroxybutyric acid, have been directly linked to RCD's (Bennett *et al*., 1994). Finally, the impaired mitochondrial β-oxidation and inhibition of the ETC and resultant mitochondrial dysfunctional due to ROS, may also lead to the formation of hepatotoxicity or liver damage (Pessayre *et al*., 1999). Elevated concentrations of the amino acids, phenylalanine and tyrosine, are commonly observed in cases of liver damage (Tessari *et al*., 2008) in addition to their metabolites: 4-hydroxyphenylacetic acid, 4-hydroxyphenyllactic acid and o-hydroxyphenylacetic acid, previously reported in the RCD patients described by Reinecke *et al*. (2012) and elevated in the treatment failure group in our investigation. This also relates to the perturbed NADH/NAD ratio, and may also be an indication of liver damage in the treatment failure group, induced either by the ROS, or other societal factors including alcohol intake, for which this Western Cape community are well known for (May *et al*., 2005; Shisana *et al*., 2005).

Oxidative stress also affects the peroxisomes in the human host. It has been previously observed, in various disease states, that the ability of the peroxisome to maintain the balance of H_2O_2-generating and H_2O_2-degradading activity, is greatly compromised. It is therefore critical to maintain peroxisomal redox balance, however, this can be compromised during conditions of severe oxidative stress, which results in peroxisomal dysfunction, or even in their degradation (Lvashchenko *et al*., 2010). Refsum's disease (RD), is an autosomal recessive inborn error of lipid metabolism and is classified as a peroxisomal disorder (Wierzbicki *et al*., 2003). The defective enzyme in this disease state is phytanoyl-coenzyme A hydroxylase, which causes impaired α-oxidation of 3-methyl-branched fatty acids, including phytanic acid, and the subsequently accumulation of these fatty acids and excretion in the urine (Willis *et al*., 2001). The accumulating phytanic acid is catabolised by

means of ω-oxidation to eventually produce a variety of 3-methyl-organic acids, including 3-methyladipic acid (Wierzbicki *et al*., 2003). As mentioned previously, 3-methyladipic acid was elevated in the urine of the treatment failure group, indicative of dysfunctional peroxisomes. This is further supported by the increased pipecolic acid detected in these patients. Pipecolic acid is an intermediate of lysine metabolism, and synthesised by the peroxisomal enzyme, L-pipecolate oxidase. A loss in activity of this enzyme, results in an increase of pipecolic acid. Furthermore, elevated levels of pipecolic acid have also been reported in other peroxisomal disorders, including Refsum disease (Gould *et al.*, 2001; Willis *et al*., 2001).

Hypochlorous acid (HOCl), also classified as reactive oxygen species, generated by the mycloperoxidase enzyme in neutrophils, is produced during a variety of diseases caused by infectious bacteria (Halliwell & Aruoma, 1991). When the human body is in a state of oxidative stress, a wide range of antioxidant genes are regulated, in order to protect cells from the damaging effects of ROS. Studies suggest that uric acid can also act as an antioxidant *in vivo*. Uric acid is oxidized, by HOCl, to produce allantion, oxaluric and parabanic acid (Kaur & Halliwell, 1990). Parabanic acid was elevated in the urine of the failed treatment outcome group, suggesting that a greater amount of ROS are possibly generated in these patients, possibly due to an underlying RCD in these patients.

It is also well known, that DNA damage, generated by oxidative stress, has been implicated in the pathogenesis of various diseases (Cooke *et al*., 2008). DNA damage can be induced by the activation of nucleases or direct reaction of ROS with the DNA. As a result of this interaction, the DNA is modified and subsequently damaged (Halliwell & Aruoma, 1991). Oxidatively modified DNA lesions, have also been reported to be found in urine (Cooke *et al*., 2008). 2-Deoxyribonic acid is derived from deoxyribonate, which is produced as part of the bistranded lesions in response to oxidative stress. 2-Deoxyribonic acid was also detected in elevated concentrations in the treatment failure group, which is further indicative of elevated oxidative stress in this group, again supporting the notion of a potentially underlying RCD.

5.2.1.2.3 Markers of exogenous origins

Elevated levels of furan compounds, including: 5-hydroxymethylfuroic acid, furan-2,5-dicarboxlic acid and 2-furoylglycine, were also observed to be elevated in the treatment failure group. These metabolites could be a result of exposure to fural or furyl alcohol (Flek & Sedivec, 1976), which is highly unlikely in this population. It is also possible that these metabolites could have origins from furan derivatives found in food sources such as coffee, cocoa or various cooked meat products (Pettersen & Jellum, 1972). The food-borne mold, *Aspergillus sp,* which is able to live and reproduce in the gastrointestinal tract of the human body, and also known to produce these furan compounds detected in the treatment failure group (Sumiki, 1929; Sumiki, 1931; Kawarda *et al*., 1955). 2-Furoylglycine, is believed to be produced as a detoxification product of furan-2,5-dicarboxylic acid (Mrochek & Rainey, 1972). Reduced concentrations of these furan compounds have also been observed in the urine of autistic children, after treatment with nystatin, an indication that these compounds are of yeast and/or fungus origin. Additionally oxalic acid, also observed in elevated concentrations in the treatment failure group, is known to be produced as a major organic acid from various species of *Aspergillus* (Alam *et al*., 2002). Oxalic acid is also known to be an end product of ascorbic acid oxidation and is also elevated in the urine of individuals who ingest ascorbic acid (Robitaille *et al*., 2009). Considering this, the increase of these three furan compounds and oxalic acid, in the urine of the treatment failure group, is most likely due to microbial overgrowth of yeast and/or fungus species such as *Aspergillus* in their gastrointestinal tract.

Furthermore, elevated urinary excretion of phenylacetic acid, 3-(3-hydroxyphenyl)-3-hydroxypropapoinc acid (HPHPA) and benzoic acid, are also associated with bacterial overgrowth in the intestinal tract (Lord & Brally, 2008). Cases of bacterial overgrowth have also been reported to be caused by diverse problems, which include tuberculosis (Singh & Toskes, 2003). Phenylacetic acid, a derivative of phenylalanine, may also be produced from dietary sources (phenylalanine) by means of gut micriobiota, and also found to be elevated in the urine of patients suffering from phenylketonuria. It is also known that elevated amounts of HPHPA are excreted in the urine of patients suffering from autism and schizophrenia (Shaw, 2010; Persico & Napolioni, 2012). Benzoic acid was identified as one of the first compounds found to be increased in the urine of individuals with intestinal bacterial overgrowth. Benzoic acid is produced by the bacterial catabolism of dietary polyphenols and normally conjugates with glycine to form hippuric acid in the liver (Lord &

Brally, 2008). The release of elevated concentrations of phenylacetic acid, HPHPA and benzoic acid, as observed in the urine of the treatment failure group, further supports the notion that there is an increased disturbance of the normal microbial ecosystem in these patients, due to a possibly greater sensitivity to infection of *M. tuberculosis* in these patients.

Polyphenols are naturally occurring compounds in plants, especially fruits, vegetables and some beverages, including wine, coffee and tea. Dietary intake of these polyphenols has shown to contribute to human health and have various beneficial effects (Gonthier *et al*., 2002). Phenolic compounds, including quercetin (Tian *et al*., 2009), catechins (Morita *et al*., 2003), caftaric acid and coutaric acid (esters of caffeic and *p*-coumaric acid with tartaric acid)(Gonthier *et al*., 2002) are found to be abundant in wine. Four metabolites derived from polyphenols found in wine, were observed in the urine of the TB patients, and include: dihydroferulic acid, vanillic acid, 3-hydroxyhippuric acid and 4-methylcatechol. Once wine is ingested, quercetin is degredated by interstinal microorganisms to produce 4-methylcatechol. Additionally caffeic acid and *p*-coumaric acid are produced by means of microbial esterase activity. In the liver caffeic acid is methylated to form ferulic acid, which is further metabolized to produce vanillic acid and dihydroferulic acid (Booth et *al*., 1957; Gumbinger *et al*., 1993; Rechner *et al*., 2002). Microflora in the intestinal tract also transform cinnamic acid into 3-hydroxphenylproionic acid and 3-hydroxyhippuric acid (Gonthier *et al*., 2002).The four metabolites mentioned above, are showed to be elevated in the urine of the treatment failure group. Due to the fact that these patients live in the rural areas of the Western Cape, where it is known that large quantities of wine are consumed recreationally, the assumption can be made that these compounds found to be present in the urine, are derived from increased wine consumption in this group, accompanied by the RCD's and microflora disruptions. The significant lower levels of the above mentioned compounds in the cured patients may suggest that excessive alcohol consumption could directly or indirectly (poor compliance to treatment) contribute to treatment failure.

5.2.2 Metabolomics comparisons of successful and unsuccessful treatment outcomes during treatment

Prior to univariate and multivariate statistical analysis, a variety of data clean-up steps were applied, as described in section 4.5.3, to the data set compiled by subtracting the metabolite

concentrations at week 4 of treatment from week zero (before treatment). Prior to multivariate and univariate statistical analysis, a log transformation function was applied. PCA, PLSDA, paired *t*-test and effect size calculations, were subsequently applied to the processed dataset containing 231 aligned compounds, 126 of which could be annotated by comparison of their mass spectra and retention times to that of libraries compiled from previously injected standards.

5.2.2.1 PCA differentiation between successful and unsuccessful treatment outcome groups during treatment

A PCA was performed on the GCxGC-TOFMS generated dataset in order to determine whether or not a natural separation exists between the cured and failed treatment group, while both are on anti-TB treatment. Using the 231 variables in the processed data set described above, a clear differentiation of the two groups was achieved by the first two PCs of the PCA (Figure 5.7). The total amount of variance explained by the first three PCs (R^2X cum) was 56.5%, of which PC 1 explained 42.7%, PC 2 explained 9.2% and PC 3 explained 4.6%. This clear differentiation achieved, can be ascribed to the differences in the extracted metabolite profiles of the individual patients in each group. This result indicates that these two outcome groups respond differently to the anti-TB drugs administrated during the six month treatment regimen. It may also be attributed to the altered host metabolome caused by the *M. tuberculosis* infection and seen in the treatment failure groups as described in section 5.3.1.2.2, as this factor is still relevant and present in this comparison.

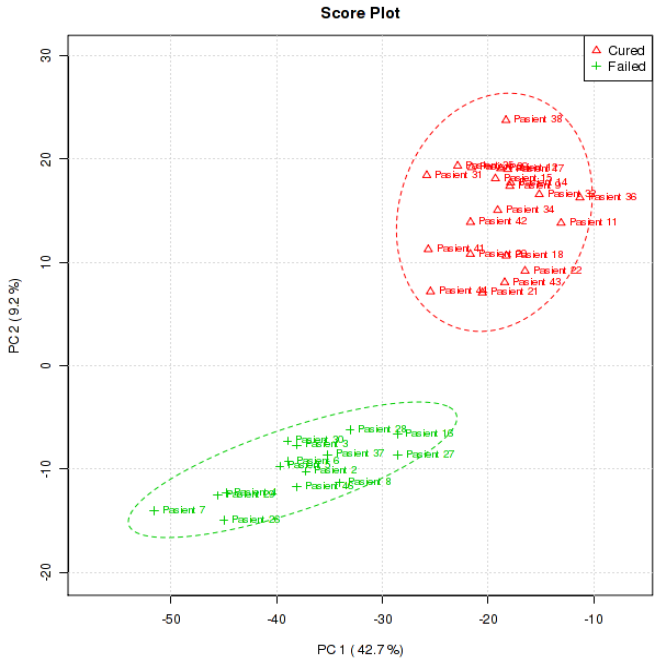


Figure 5.6: PCA scores plot illustrating PC1 vs. PC2 using data calculated by determining the difference in metabolite concentrations at week 4 of treatment and week zero (before treatment), subsequently to organic acid extraction and GCxGC/TOFMS analysis, indicating a differentiation of the two outcome groups. The variance explained by each PC is given in parenthesis.

5.2.2.2 Metabolite marker identification

In our quest to identify those metabolite markers better describing the variation between the cured and failed treatment groups, during the treatment phase, the results from the univariate and multivariate statistical methods, were once again used to compile a list of metabolite markers best describing the variation seen between these groups at these time points. As illustrated in Figure 5.6, a natural differentiation between the two groups exists and all the metabolites associated or responsible for this differentiation were subsequently identified. All metabolites with a PCA modelling power > 0.5 were selected (Brereton, 2003), representing those metabolites/variables best describing the variation seen in the PCA projection. Furthermore, PLS-DA was used to determine the VIP values of all the metabolites present in the two sample groups. Metabolites with a VIP value > 1 were selected as potential markers (Chong & Jun, 2005) and represent those metabolites which vary the most between the compared sample groups. This PLS-DA model used had a

modelling parameter, R_2Y (cum) of 97.5%, indicative of the total explained variation of the response, Y. The cross-validated variation, Q_2 (cum) explained by the response Y, was 74.4%.

A number of univariate statistical methods was also used from further metabolite selection, and included effect size and paired *t*-test calculation. Metabolites with an effect size > 0.5 and/or and paired *t*-test p-value < 0.05 were considered as being most significant. A detailed explanation for these statistical methods is given in Chapter 3, section 4.7. Using the above statistical methods, the final list of metabolite markers were compiled from a combination of the selected PCA and PLS-DA markers, having an ES > 0.5 (Figure 5.7). The *p*-values for each metabolite marker is also reported (Table 5.4).

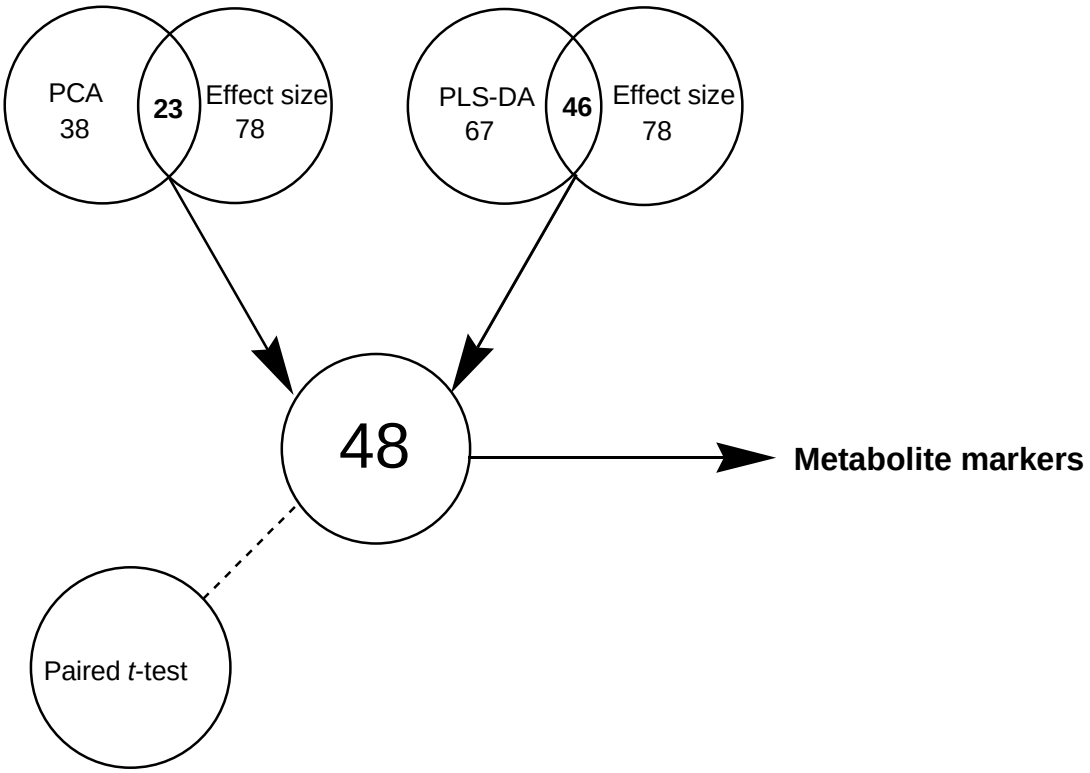


Figure 5.7: Statistical approach used to identify metabolite markers which best explain the variation between the successful and failed outcome during treatment.

Of the 48 compounds selected as metabolite markers using this above selection approach, 20 were annotated using libraries prepared from injected standards by comparison of their mass spectra and retention times. The PCA power, PLS-DA VIP, average concentration (difference between week 0 and week 4), standard deviation, effect size and paired *t*-test (*p*-value) of each of the 20 metabolites are provided in Table 5.4. These metabolites will subsequently be used in order to explain the differences that exist between the two outcome groups, during the use of the standard TB treatment regimen.

Table 5.4: Metabolite markers best describing the variation between the cured and failed treatment oucome groups during the standard TB treatment regimen.

Metabolite Name			**Successful treatment outcome**	**Failed Treatment outcome**		
	PCA Power	**PLSDA VIP**	***Average concentration (w4-w0) ± standard deviation (mg/g creatinine)**	***Average concentration (w4-w0)± standard deviation (mg/g creatinine)**	**Effect Sizes**	**Paired *t*-test (p-value)**
Markers associated with an altered host metabolism and further aggravated by anti-TB treatment						
Glycolic acid	0.661	2.054	0.796 ± 37.512	3.217 ± 17.060	0.686	<0.001
o-Hydroxyphenylacetic acid	0.637	0.913	0.017 ± 2.891	0.435 ± 3.175	0.606	0.065
Cis-Aconitic acid	0.552	1.505	-2.888 ± 56.124	-14.767 ±61.097	0.602	0.015
2,4-Dihydroxybutyric acid	0.540	0.972	4.931 ± 25.299	-19.262 ± 6.489	0.679	0.052
Homovanillic acid	0.519	0.879	-0.936 ± 4.607	1.165 ± 10.844	0.599	0.029
p-Hydroxymandelic acid	0.486	1.261	-1.115 ± 8.491	0.408 ± 7.415	0.558	0.057
Marker associated with anti-TB drug induced side effects						
3-Hydroxyisobutyric acid	0.489	1027	-2.219 ± 7.614	5.583 ± 20.596	0.777	<0.001
2-Ethylhydracrylic acid	0.467	1.948	-2.681 ± 16.522	2.82 ± 8.685	0.578	0.062
Hexanoic acid	0.455	1.893	-0.767 ± 3.412	0.122 ± 0.910	0.613	<0.001
Markers of exogenous origin						
Ascorbic acid	0.757	2.884	3.244 ± 2.559	-1.942 ± 1.678	0.530	<0.001

Phenoxyacetic acid	0.681	1.788	-0.194 ± 0.771	0.182 ± 1873	2.138	<0.001
Acetylsalicylic acid	0.679	2.855	0.011 ± 0.066	8.453 ± 31.029	1.779	<0.001
3-(3-hydroxyphenyl)-3-hydroxypropanoic acid	0.536	1.669	-2.853 ± 71.970	-8.239 ± 80.455	0.769	0.011
Ethyl vanillic acid	0.527	1.853	0.044 ± 0.208	4.310 ± 15.178	1.953	<0.001
Vanillic acid	0.482	1.299	0.176 ± 7.771	1.441 ± 9.433	0.829	0.037
Dihydrofurilic acid	0.451	1.281	-0.012 ± 0.376	0.314 ± 0.827	0.690	0.063
Cis-4-hydroxycyclohexane carboxylic acid	0.443	1.799	-0.287 ± 2.732	-0.623 ± 0.904	0.815	0.003
Unclassified						
N-Formyl-glycine	0.538	1271	1.546 ± 3.066	-1.812 ± 1.535	0.612	0.059

* Average concentration = metabolite concentrations at week 4 (w4) - metabolite concentration at week 0 (w0 - prior to when treatment commenced).

The markers identified could be grouped according to those originating from: 1) An altered host response to TB and further aggravated by anti-TB medication, 2) anti-TB drug induced side effects, and 3) exogenous origins.

5.2.2.2.1 Markers associated with an altered host response to TB and further aggravated by anti-TB drugs

As described in section 5.2.1.2.2, a large number of metabolites were found to be elevated in the urine of the treatment failure group, when investigating the metabolite profiles of these patients collected at time of diagnosis, before treatment commenced. The detailed interpretation of theses markers, unanimously suggested that those individuals in the treatment failure group, possibly have an underlying RCD or mitochondrial dysfunction, which results in a far greater severity to the infection, and subsequently, a reduced capacity for a successful treatment outcome, also partly due to the fact that the mitochondria plays a critical role in anti-TB drug metabolism. It was also clear that *M. tuberculosis* infection

results in a host macrophage response, altering the host metabolism. Results indicated, in the treatment failure group, that the TB infection is accompanied by a more pronounced ROS production and severity of oxidative stress, promoting inflammation and DNA damage, and ultimately resulting in ETC inhibition, decreased mitochondrial function, reduced ATP production, and TCA cycle inhibition. This may further lead to impaired mitochondrial fatty acid β-oxidation in this group, promoting dicarboxylic acid production via ω-oxidation. Additionally, the more pronounced oxidative stress and the consequential increased H_2O_2 production, may cause peroxisomal dysfunction or degradation, and thus the consequential impaired α-oxidation, resulting in increased 3-methyl-branched chain fatty acid concentrations in the urine. Further support for the possible underlying RCD in the treatment failure group, were the metabolite markers related to elevated levels of various neurotransmitters (such as GABA, dopamine and epinephrine) and amino acid metabolism (especially tryptophan, phenylalanine, tyrosine and isoleucine). Furthermore, in the treatment failure group were metabolites indicating a possible pronounced hepatotoxicity or liver damage, in addition to markers implicating a tendency towards elevated blood glucose levels, and diabetes mellitus.

Considering the results in this section, once again a number of markers were identified in the treatment failure groups which can be linked to a varied alteration to host metabolism due to infection, and which may be further aggravated due to the administration of the anti-TB treatment. Cis-aconitic acid for instance, seen to be dramatically decreased in the treatment failure group once the medication was given. Cis-aconitic acid is an intermediate of the Krebs cycle produced by dehydration of citric acid, and facilitated by means of the enzyme cis-aconitase. The enzyme requires a sulfhydryl group from cysteine or glutathione for optimal functioning and is activated by ferrous ions. Thus dysfunctional cysteine and methionine metabolism as well as depletion of glutathione, commonly observed in cases of increased oxidative stress, may cause sub-optimal enzyme activity of cis-aconitase (Kovacevic *et al*., 1970). As a result less cis-aconitic acid will be produced. A decrease of urinary cis-aconitic acid has been observed in the urine of the treatment failure group, during treatment with anti-TB drugs. ROS generated by administration of anti-TB drugs have been known to impair the function of ETC (Loots *et al*., 2004) and which may subsequently lead to impaired Krebs cycle function. Oxidative stress induced by the generated ROS may also have led to glutathione depletion, which my caused low enzyme activity of cis-asonitase and resulted in low levels of cis-aconitic acid observed in the urine.

As previously discussed, reduced mitochondrial production of ATP (due to mitochondrial dysfunction) can be directly compensated for by the upregulation of carbohydrate, protein and triacylglycerol catabolism, by the action of various neurotransmitters such as dopamine (Reinecke *et al*., 2012). Homovanillic acid was detected in elevated concentrations in the treatment failure group at time of diagnosis, and is indicative of elevated levels of the neurotransmitters dopamine (Frankenhaeuser *et al*., 1986). Homovanillic acid was once again detected as a marker in the urine of the treatment failure group during anti-TB drug treatment phase in these patients. This increase is possibly due to the underlying RCD, probably worsened by the additional oxidative stress induced by the anti-TB drugs (Loots *et al*., 2004). Additionally, elevated 4-hydroxybutyric acid was observed in the urine of the treatment failure group at time of diagnosis, and was associated with an increase GABA metabolism (Sah & Schwartz-Bloom, 1999). Additionally, inhibition of the ETC and Krebs cycle, would most likely also result in the accumulations of succinic semialdehyde, consequently resulting in the synthesis of 4-hydroxybutyric acid (Pearl *et al*., 2011). 4-Hydroxybutyric acid, as well as 2,4-dihydroxybutyric acid (a metabolite produced by α-oxidation of 4-hydroxybutyric acid), are also previously reported to be elevated in the urine of patients suffering from a succinic semialdehyde dehydrogenase deficiency (Brown *et al*., 1987). In this investigation, 2,4-dihydroxybutyric acid was comparatively reduced in the treatment failure group, during treatment phase of this experiment. This decrease may be attributed to impaired α-oxidation, due to the previously described inhibition of the peroxisomes due to RCD, being pronounciated by the effects of the anti-TB medication. It is thus possible that the oxidative stress caused by the anti-TB drugs contributes to the existing peroxisomal dysfunction assumed to be present in the treatment failure group, due to oxidative stress and the impact this has on the peroxisomes (Lvashchenko *et al*., 2010). Further confirmation of this was the elevated levels of glycolic acid detected (van Woerden *et al*. 2006).

And lastly, once again, the markers for liver damage, o-hydroxyphenylacetic acid and p-hydroxymandelic acid (Wadman *et al*., 1971), were detected in the treatment failure group, after the administration of anti-TB drugs. It is well known, that these drugs are associated with a wide variety of adverse side effects, including hepatotoxicity. Considering this, the impaired β-oxidation and inhibition of ETC, as a result of underlying mitochondrial dysfunction in this group due to the RCD, would further be aggravated by the ROS produced by the anti-TB drugs, dramatically contributing to the liver damage in this patient group (Pessayre *et al*., 1999).

5.2.2.2.2 Marker associated with anti-TB drug induced side effects

Loots *et al*, (2004), used a metabolomics research approach to study the possible effects of a combined anti-TB drug known as Rifater, by studying the changes induced by the drug on the urine organic acid and free radical profiles of rats. They found that the Rifater treatment increased the hydroxyl radicals and various organic acids characteristic of a multiple acyl-CoA dehydrogenase defect (MADD), also known as glutaric aciduria type II. They subsequently suggested that Rifater treatment induces its associated side effects by inhibition of either the electron transport flavoprotein (ETF) or the ETF dehyrogenase enzyme (Loots *et al*, 2004). Hexanoic acid, 3-hydroxyisobutyric acid, 2-ethylhydracrylic acid and 3-methylglutaconic acid have all been found to be excreted in the urine of patients suffering from glutaric aciduria type II (Prezyrembel *et al*., 1976; Reineke *et al*., 2012; Wortmann, 2012). Elevated concentrations of 3 of these metabolites were also observed in the urine of the treatment failure group, while on anti-TB medication. It is therefore proposed that the anti-TB drugs administered by these patients gave rise to the metabolite profile observed, and the treatment failure group are more susceptible to these disturbances, probably due to the underlying RCD previously detected. Furthermore, these alterations have also been linked with the side effects associated with the administration of the anti-TB drugs (Loots *et al*., 2004), and hence, it is also possible, the treatment failure group may be more susceptible to, or may experience these side effects more severely.

.

5.2.2.2.3 Markers of exogenous origin

A variety of metabolites associated with medications or treatments other than that of the anti-TB drugs administrated during this treatment period were also observed in the urine of these TB patients. Ethyl vanillic acid (3-ethoxy-4-hydroxybenzoic acid) is produced and excreted in the urine once ethyl vanillin is consumed. Ethyl vanillin is an important food additive and flavor enhancer and is found in a wide variety of food product (Ni *et al*., 2005), and is also used as an artificial flavoring agent in medication, including valproic acid (Mamer *et al*., 1985). Elevated concentrations of ethyl vanillic acid were observed in the urine of the treatment failure group, and may be attributed to additional medication which is administrated, possibly to relive symptoms of anti-TB drug induced side effects.

Phenoxyacetic acid is classified as a diuretic agent, which is chemically different from other types of diuretic agents (Paul *et al*., 1965). Diuretic agents are used to treat a number of conditions, such as hypertension, hypercalcemia, hypercalciuria and acute renal failure (Krumlovsky & del Greco, 1976). One of the side effects commonly associated with isoniazid administration is interstitial nephritis (inflammation of the nephrons in the kidneys), which may lead to renal failure (Arbex *et al*., 2009).

When acetylsalicylic acid (asprin) is ingested, it is first metabolised to produce salicylic acid which in turn in converted to 2,5-dihydroxybenzoic acid (gentisic acid) (Zaugg *et al*., 2001). Acetylsalicylic acid and its related metabolite are found to be excreted in the urine of individuals who consume aspirin (Cham *et al*., 1982). As previously mentioned in section 2.3, administration of anti-TB drugs are associated with a wide variety of side effects which include hyperuricemia, arthralgia, abdominal pain and headaches (Arbex *et al*., 2009). It is therefore possible that the increase in concentration of acetylsalicylic observed in the urine of the patients in the treatment failure group is due to aspirin ingestion to relieve these side effects. We previously hypothesized that these patients in the treatment failure group, might have more anti-TB drug associated side effects and they probably take aspirin in order to relieve the pain caused by these side effects.

An increased concentration of urinary ascorbic acid (vitamin C) was observed in the successful treatment outcome group. Ascorbic acid is a water soluble vitamin and an essential nutrient in the human diet. This vitamin is found in fruits, vegetables and berries, and improves immune system function and also contributes to the prevention in the spread of infectious agents. It has also been found that administration of ascorbic acid by patients diagnosed with pulmonary TB, have shown to accelerate the healing of lung fibrosis and other damages caused by this infection (Bakaev & Duntau, 2004). An increase in urinary ascorbic acid was detected in the urine of the successful treatment outcome group may be due to these patients additionally taking ascorbic acid during the treatment regimen, in order to help improve their immune system and overall health, and hence contributing to a successful outcome. This is an interesting observation which should be investigated further.

A variety of polyphenols, including: vanillic acid and dihydrofurilic acid were identified in the urine of the treatment failure group at time of diagnosis (section 5.2.1.2.3). These

polyphenols and their derivatives are assumed to originate from wine consumption, due to the fact that these patients live in the rural areas of the Western Cape, where it is known that large quantities of wine are consumed recreationally. Vanillic acid and dihydrofurilic acid were again found to be elevated in the urine of the treatment failure group. Thus the possibility exists that these patients might have continued to consume large amounts of wine during the treatment regimen. Factors such as alcohol consumption during treatment regimens have been implicated as a contributor of non-adherence and result in an increased risk of TB treatment failure (Jaiswal *et al*., 2003).

As previously described in section 5.2.1.2.3, various metabolites associated with bacterial overgrowth, were found to be elevated in the urine of the treatment failure group at time of diagnosis. It was assumed that an increase disturbance of the normal microbial ecosystem were present in these patients, due to a possibly greater sensitivity to infection of *M. tuberculosis* in these patients. 3-(3-Hydroxyphenyl)-3-hydroxypropapoinc acid (HPHPA) (also identified at time of diagnosis in treatment failure group), and cis-4-hydroxycyclohexanecarboxylic acid, both associated with bacterial overgrowth in the gastrointestinal tract (Lord & Brally, 2008; Kronick *et al*., 1983), were all found to be decreased in the urine of all patients, and especially those of the treatment failure group, during the treatment phase. This decrease observed may be due to administration of the potent anti-TB drugs, which possibly eliminated the aggressive bacterial species in the gastrointestinal tract, causing disturbance of normal microbial ecosystem.

5.2.3 Metabolomics comparisons of successful and unsuccessful treatment outcomes over time

Prior to univariate and multivariate statistical analysis, a variety of prepossessing steps were applied to the data as described in section 4.6.3. For the application of both multivariate and univariate statistical analysis, a log transformation function was applied. Time dependant statistical analysis which included ANOVA (univartiate) and ASCA (multivariate), were subsequently applied to the processed dataset. This dataset contained the processed data of each TB patient at all-time points of which urine samples were collected, including time of diagnosis and weeks 1, 2, 4 and 26 of anti-TB treatment regimen. This dataset contained

231 aligned compounds, 126 of which could be annotated by comparison of the mass spectra and retention times of each compound, to that of libraries compiled from previously injected standards.

5.2.3.1 ANOVA and ASCA statistical analysis between successful and unsuccessful treatment outcome groups

Two-way ANOVA and ASCA were performed on the GCxGC-TOFMS generated dataset in order to determine important metabolites associated with the dominant variation between the successfully treated group and the unsuccessfully treatment group over time. We used time dependent statistical analysis as a confirmation of the variables we identified and discussed at time of diagnosis (section 5.2.1.2) and during the treatment period (section 5.2.2.2). Furthermore, we were also hoping to identify additional metabolites which differ significantly between the two outcome groups over time. Both two-way ANOVA and ASCA provided 2 lists of important variables, including metabolites which differ the most in accordance to outcome, as well as metabolites associated with the interaction of time and outcome. Only variables associated with outcome and the interaction between time and outcome were used, due to the fact that variables which significantly change over time, independently of treatment outcome, were not seen to be relevant to our study. As illustrated in Figure 5.7 a and b, it is evident that differences exist between the two outcome groups at all time points, however to a lesser extend at week 4 of treatment.

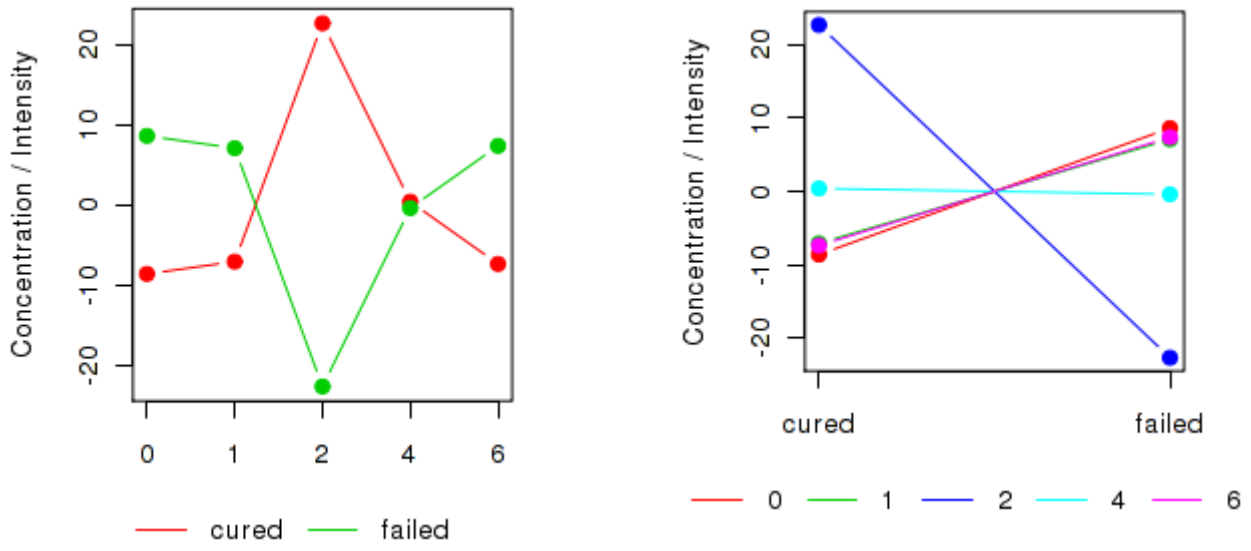


Figure 5.8: Major patterns associated with the interaction between time (0, 1, 2, 4, 6 representing week 0, 1, 2, 4 and 26) and treatment outcome (Indicated as cured and failed) as identified by ASCA analysis (figures generated on MetaboAnalyst).

In our quest to identify metabolite markers, results from ASCA and ANOVA were combined in order to compile a list of metabolite markers which have shown to be most significant between the two treatment outcome groups. In the case of ANOVA, all outcome and interaction metabolites with a $p \leq 0.05$ were selected, representing those metabolites/variables best describing the variation of the two outcome groups over the duration time of anti-TB treatment. Furthermore, ASCA was used to determine the Leverage and SPE value of all the variables present in the two outcome groups. Variables with a Leverage value between 0 and 1 (Nueda *et al*., 2010), were selected, as they represent well-modelled variables and which vary most between the compared sample groups over time.

5.2.3.2 Metabolite marker identification

Using the above statistical methods, the final list of metabolite markers were compiled from a combination of all the selected ANOVA and ASCA variables associated with treatment outcome and the interaction between time and treatment outcome, which have shown to be most significant. Of the 45 compounds selected as metabolite markers using this above selection approach, 35 were annotated, using libraries prepared from injected standards by comparison of their mass spectra and retention times. Of these 35 metabolites, 21 were

previously identified at time of diagnosis and/or during treatment. ANOVA p-values for time, outcome and interaction, as well as the ASCA leverage and SPE values of each of the 35 metabolites are provided in table 5.5. These metabolites will subsequently be used in order to explain the differences that exist between the two outcome groups over time.

Table 5.5: Metabolite markers best describing the variation between the successful and faild treatment oucome groups from time of diagnosis till the end of the six month treatment regimen

	ANOVA p-value			ASCA	
Metabolite name	**Time**	**Outcome**	**Interaction**	***Leverage**	***SPE**
Markers assosiated with an altered host metabolism					
Quinolinic acid	0.123	0.021	<0.001	NA	NA
Pyroglutamic acid	0.008	<0.001	0.021	NA	NA
3,4-Dihydroxybutyric acid	0.229	0.009	0.039	NA	NA
2,4-dihydroxybutyric acid	0.144	0.022	0.167	NA	NA
Homovallinic acid	0.212	0.046	0.122	NA	NA
Glucuronic acid	0.118	0.082	<0.001	NA	NA
2-Methyl-3-hydroxybutyric acid	0.219	0.023	0.03	NA	NA
3-Hydroxyvaleric acid	0.412	0.86	0.04	NA	NA
Methylmalonic acid	0.303	0.862	0.031	NA	NA
3-Hydroxysebacic acid	0.424	0.029	0.257	NA	NA
Seberic acid	0.024	0.016	0.208	NA	NA
Azelaic acid	0.117	0.021	0.144	NA	NA
Pimilic acid	0.029	0.027	0.05	NA	NA
1-Decene	0.03	<0.001	<0.001	NA	NA
Octane	0.308	<0.001	0.067	NA	NA
Pyruvic acid	0.861	0.019	0.795	NA	NA
Malic acid	0.550	0.194	0.031	NA	NA
p-Hydroxymandelic acid	0.163	0.029	0.168	NA	NA
4-Hydroxybenzoic acid	0.938	0.921	<0.001	NA	NA
Markers associated with anti-TB drug treatment					
Hexanoic acid	0.103	0.024	0.052	NA	NA
2-hydroxybutyric acid	0.246	0.765	0.005	NA	NA
Isobutyric	0.153	0.039	0.138	NA	NA
3-Hydroxyisobutyric acid	0.188	0.867	0.011	NA	NA
Markers of exogenous origins					
Ethyl vanillic acid	0.001	<0.001	<0.001	0.048	9.815
2,4-Dihydroxybenzoic acid	0.091	0.681	0.022	NA	NA
Phenoxyacetic acid	<0.001	<0.001	<0.001	0.093	4.068
Acetylsalicylic acid	0.003	<0.001	<0.001	0.067	5.634
Salicylic acid	0.11	0.008	0.005	NA	NA
Ascorbic acid	0.119	0.015	0.031	NA	NA
4-Methylcatechol	0.987	0.432	0.01	NA	NA
Vanillic acid	0.226	0.011	0.308	NA	NA

3,4-Dihydroxycinnamic acid	0.074	0.082	0.012	NA	NA
Hippuric acid	0.308	0.016	<0.001	NA	NA
Cis-4-hydroxycyclohexane carboxylic acid	0.117	0.002	<0.001	NA	NA
3-(3-hydroxyphenyl)-3-hydroxypropanoic acid	0.052	0.026	0.004	NA	NA
Unclassified					
N-formyl-glycine	0.842	<0.001	0.122	NA	NA

*Ethyl vanillic acid, phenoxyacetic acid and acetylsalicylic acid is the only three known compounds Identified by ASCA, and are therefore the only compounds with leverage and SPE values. (NA - not applicable)

5.2.3.2.1 Markers associated with an altered host response to TB and further aggravated by anti-TB drugs

As described in section 5.2.1.2.1, an increase of quinolinic acid was observed in the treatment failure group at time of diagnosis and ascribed to result from the infection with *M. tuberculosis,* resulting in a host macrophage response, engulfing the invading bacteria and subsequently tryptophan catabolism to quinolic acid is increased (Blumenthal *et al*., 2012). The Two-way ANOVA analysis confirmed the identification of quinolinic acid as an important interaction marker (between treatment outcome and time), confirming that the greatest difference in quinolinic acid concentration between the two outcome groups exists at time of diagnosis (Appendix 2a). Illustrated in Appendix 2a, an increase in quinolinic acid is observed during treatment. Shibata *et al*, (2001) have found that administration of dietary pyrazinamide (one of the anti-TB drugs administrated by the patients in this study) increases the metabolism of tryptophan to niacin, as well as resulting in elevated levels of various tryptophan metabolites, including quinolinic acid, excreted in the urine. Thus the increase of quinolinic acid observed during treatment may be attributed to the anti-TB drug pyrazinamide.

Pyroglutamic acid is a metabolite related to glutathione, and is produced in the γ-glutamyl cycle (GGC), which is active when a high demand for glutathione is needed, including states

of increased oxidative stress (Lord & Bralley, 2008). Reineke *et al*. (2012) previously observed elevated concentrations of pyroglutamic acid in the urine of patients suffering from RCD's. This increase was attributed to interrelated involvement of the NAD redox state, the formation of ROS and the glutathione redox state, observed in RSDs. Pyroglutamic acid was also identified as an important metabolite of interaction between the two outcome groups, and it is illustrated (Appendix 2b), the greatest variation of this metabolite exists at time of diagnosis and at week 26, with elevated amounts detected in the treatment failure group at these time points. This elevation may be attributed to the alterations made in the host metabolism by *M. tuberculosis* infection as previously reported in section 5.2.1.2.2. which is still present in these patients at week 26.

3,4-Dihydroxybutyric acid and 2,4-dihydroxybutyric acid, are both metabolites of 4-hydroxybutyric acid, which was associated with elevated levels of GABA, and accumulation of succinic semialdehyde. 3,4-Dihydroxybutyric acid was found to be elevated in the urine of the treatment failure group at time of diagnosis, which was also confirmed by the ANOVA analysis in this section (Appendix 2c). Subsequently, a decrease in 2,4-dihydroxybutyric acid, an α-oxidation product of 4-hydroxybutyric acid, was detected in the treatment failure group, during the treatment regimen (section 5.2.2.2.1), due to an impaired α-oxidation caused by the anti-TB drugs. The ANOVA analysis confirmed the 2,4-dihydroxybutyric acid, as a significant marker of outcome and confirmed that the greatest effect occurs during treatment (Appendix 2d).

As previously discussed, reduced mitochondrial production of ATP (due to mitochondrial dysfunction) can be compensated for by the upregulation of carbohydrate, protein and triacylglycerol catabolism, through the action of various neurotransmitters such as dopamine, and hormones such as adrenaline (Reinecke *et al*., 2012). Homovanillic acid was detected in elevated concentrations in the treatment failure group at time of diagnosis, and was indicative of elevated levels of dopamine (Frankenhaeuser *et al*., 1986). Glucuronic acid was also detected in elevated concentrations in the urine of the treatment failure group, and associated with increased blood glucose levels (Lawson *et al*., 1976), which is most likely due to elevated adrenalin levels, or severe TB disease causing insulin resistance (Faurholt-Jepsen *et al*., 2011). It was indicated by time dependant statistical analyses that these two metabolites differed most between the two outcome groups at time of diagnosis and showed a general increase in the treatment failure group over time during treatment (Appendix 2e and 2f). This elevation during the treatment period may be attributed to additional oxidative

stress caused by the anti-TB drugs, resulting in a more severe case of mitochondrial dysfunction and neurotransmitter release, and consequently reduced ATP production.

Furthermore, elevated levels of 2-methyl-3-hydroxybutyric acid and 3-hydroxyvaleric acid were observed in the urine of the treatment failure group at time of diagnosis. This was assumed to be due to abnormalities occurring in isoleucine metabolism, which could be a result of the dysfunctional mitochondria in this patient group. The two isoleucine metabolites were subsequently identified as important metabolites of interaction (Table 5.4). Illustrated in Appendix 2g and 2h, the main alterations to these metabolites are at time of diagnosis, and additionally a decrease during the treatment period, followed by an increase in the urine of the treatment failure group at week 26. Additionally methylmalonic acid was also identified as an interaction metabolite by ANOVA analysis, reduced in the treatment failure group (Appendix 2i). Elevations of methylmalonic acid has been previously reported to occur in the urine of patients suffering from the autosomal recessive inborn error of metabolism, known as methylmalonic acidemia (Li *et al*., 2007). Methylmalonic acidemia is a disorder of propionic acid degradation, derived in part from isoleucine and valine catabolism (Salway, 2010), and subsequent alterations in the failed treatment outcome group, may also be associated with these abnormalities of isoleucine metabolism. Furthermore, the general decrease observed in these metabolites during treatment in both groups, may be attributed to a lowered bacterial load due to treatment, which subsequently results in a decreased effect of *M. tuberculosis* on the human host metabolism. When the treatment period is completed, these levels return to their original values in the treatment failure group, most likely due to the TB still being present, in addition to the underlying RCD.

As described earlier, at time of diagnosis, an increased formation of a wide variety of dicarboxylic acids, including 3-hydroxysebacic acid, suberic acid, azelaic acid and pimilic acid, were found to be present in urine of the treatment failure group. Impaired β-oxidation, caused by decreased NAD^+ levels, due to a dysfunctional ETC, was thought to result in the formation of these dicarboxylic acids by means of ω-oxidation. In this section, all the above mentioned metabolites were again identified as important markers of treatment outcome over time (Apendix 2j - 2m). The time dependent statistical analysis confirmed that at time of diagnosis, these metabolites had the greatest variation between the two outcome groups. Additionally, a decrease over the treatment period was observed for these dicarboxyclic acids, followed by an elevation at week 26 in the treatment failure group. Once again, this may be due to a reduced bacterial load due to the treatment, resulting in a decrease effect of

the bacteria on the host metabolism during this period, with the elevations at week 26 in the treatment failure group, ascribed to the fact that *M. tuberculosis* is still present in this patient group.

During oxidative stress, excess H_2O_2 and ROS are produced, which is well known to cause damage to various cell macromolecules including lipids, via lipid peroxidation (Cejas *et al*., 2004). Lipid peroxidation, is a process whereby lipids are oxidised to form lipid hydroperoxides, which may cause additional cell damage through a variety of by-products, produced by lipid hydroperoxide decomposition (Blair, 2001). Amongst these by-products, are various hydrocarbons, including: 1-decene and octane (Frankel, 1980; Riely & Cohen, 1974). These two hydrocarbons, were identified in this section as important metabolite markers of interaction and outcome, respectively (Table 5.4). As indicated in Appendix 2n - 2o, both 1-decene and octane showed a general increase in concentration in the treatment failure group during treatment, and may be attributed to the increased ROS levels produced by the anti-TB drugs (Loots *et al*., 2004), known to result in lipid peroxidation.

The mitochondrial pyruvate dehyrogenase complex (PDC), catalyses the conversion of pyruvate to acetyl coenzyme A, and is known as the rate limiting step in aerobic glucose oxidation, and thus, an integral part of energy metabolism in the mitochondria (Patel *et al*., 2012). The level of PDC activity, determines the balance between the use of alternative energy substrates, thus a marked reduction in enzyme activity will result in pyruvate accumulation and eventually an increase in lactic acid production (Brown *et al*., 1994). Reduced activity of PDC has also been reported in cases of mitochondrial dysfunction (Giulivi *et al*., 2010). Pyruvic acid was identified by ANOVA analysis as an important metabolite describing outcome between the two outcome groups. A general increase of pyruvic acid (Appendix 2p) over time in the treatment failure group, and then especially at week 26, may be due to mitochondrial dysfunction and impaired ETC found to be present in the treatment failure group (section 5.2.1.2.2), in addition to the inhibition of the Krebs cycle previously proposed. Furthermore, malic acid, an intermediate in the second half of the Krebs cycle, was also identified as an important variable of interaction by ANOVA analysis. As illustrated (Appendix 2q), malic acid is generally decreased in the treatment failure group, possibly due to the underlying RCD previously described in this patient group comparatively. Furthermore, there seems to be a decrease in the concentration of this metabolite during the treatment phase in both groups, which indicates that the ROS produced by the anti-TB

drugs, which is known to cause impairment of the ETC (Loots *et al*., 2004), most likely also caused inhibition of the Krebs cycle..

And finally, as previously mentioned, elevated concentrations of p-hydroxymandelic acid, an abnormal metabolite of tyrosine metabolism, typically excreted in urine in patients with liver damage (Tessari *et al*., 2008), was found in the treatment failure group prior to treatment and again during the treatment regimen (section 5.2.2.2.2). This was attributed to the drug induced side effects in the treatment failed group or due to other societal factors, including alcohol intake. The two-way ANOVA analysis in this section again identified p-hydroxymandelic acid as an important metabolite of outcome between the two patient groups and indicated that the greatest difference with regards to this metabolite is during time of treatment (Appendix 2r). Additionally 4-hydroxybenzoic acid has been identified as an important interaction marker between the two outcome groups. This phenolic acid is also a metabolite of tyrosine and is produced in extra hepatic tissues in the liver. 4-hydroxybezoic acid together with p-hydroxymandelic acid are used as biochemical control parameters in order to monitor the metabolic function of the liver. Elevated excretions of these metabolites have been reported in cases of liver cirrhosis (Liebich & Pickert, 1985). 4-Hydroxybenzoic acid has also been found to be elevated in the urine of the treatment failure group at time of treatment (Appendix 3). Thus the increase of 4-hydroxybezoic acid in the treatment failure group further support the above assumption made of the elevate presence of these two metabolites.

5.2.3.2.2 Metabolites associated with anti-TB drug induced side effects

As previously reported by Loots *et al*, (2004), the administration of anti-TB drugs, increases ROS levels, as well as the excretion of various organic acids, characteristic of a MADD metabolic profile. Four MADD metabolites, including hexanoic acid were found to be elevated in the treatment failure group during time of treatment as described in section 5.2.2.2.2. In addition to hexanoic acid, 2-hydroxybutyric acid, isobutyric acid and 3-hydroxyisobutyric acid, all metabolites associated with a MADD metabolic profile (Przyrembel *et al*., 1976; Reineke *et al*., 2012) have been additionally identified by time dependent analysis as important metabolites of outcome between the cured en failed patient groups. All these metabolites have also been found to be elevated in the urine of the

treatment failure group during treatment. As illustrated in Appendix 2t-2w, the major difference of the MADD metabolites between the two outcome groups were during the treatment phase. It may therefore be assumed that the patients in the treatment failure group, may not respond as well to anti-TB drug treatment or it is possible that more associated drug side effects are resulted in this group, most likely due to the underlying RCDs which potentially exist.

5.2.3.2.3 Markers of exogenous origins

A variety of metabolites associated with medication, other than that of the anti-TB drugs administrated during the treatment period, were also observed in the urine of these TB patients as described in section 5.2.2.2.3. Ethyl vanillic acid, which is an artificial flavouring agent in medication, phenoxyacetic acid (a diuretic agent) and acetylsalicylic acid, also known as aspirin, were all found to be elevated in the treatment failure group at time of treatment. It was proposed that these elevated concentrations of these metabolites were due to additional medication which was administrated in order to relieve symptoms associated with anti-TB drug induced side effects. Furthermore, ASCA analysis also identified these three variables as important interaction markers.

All the above mentioned metabolites have been identified by time dependent statistical analysis as metabolite markers of interaction between the two outcome groups, confirming that these metabolites show the most variation between the outcome groups during treatment (Appendix 2y-2ab). Additionally, the two-way ANOVA also identified salicylic acid as an important interaction marker. Salicylic acid is produced when aspirin is ingested and metabolised, which in turn, can be converted to 2,5-dihydroxybenzoic acid (Zaugg *et al*., 2001), which was elevated in the treatment failure group as discussed in section 5.2.2.2.3. As illustrated by the figure (Appendix 2aa), salicylic acid is also increased in the treatment failure group during the treatment phase of this experiment. The elevated concentrations of these medication metabolites observed at time of treatment in the treatment failure group, may be due to the need for taking these, due to a greater severity of the side effects to the anti-TB drug in this group. Ascorbic acid (an essential vitamin in the human diet) on the other hand, was found to be elevated in the urine of the successful treatment group during the entire treatment period and returning to normal after the patients were cured. Considering its

effects, it may be worthwhile investigating its action as a combination treatment with anti-TB medication, for increasing drug efficacy. Similar occurrences were reported for melatonin, which similarly to ascorbic acid, also has anti-oxidant effects, and has been shown to lower the drug induced side effects, and increase drug efficacy.

Furthermore, and as previously discussed in section 5.2.1.2.3, a number of phenolic acids were found to be elevated in the urine of the treatment failed group at time of diagnosis and during treatment. These polyphenols, including,vanillic acid and 4-methylcatechol, are well known to be present in wine (Rechner *et al*., 2001; Morita *et al*., 2003) and the elevated concentrations in the treatment failure group, were attributed to increased wine consumption. The ANOVA analyses, confirmed these markers with vanillic acid an important metabolite of treatment outcome and 4-methylcatechol an important metabolite of interaction, the greatest variation between the compared groups was at time of diagnosis and during treatment (Appendix 2ae and 2af) respectively. Further confirmation of this, was the elevated concentrations of 3,4-dihydroxycinnamic acid, 2,4-dihydroxybenzoic acid and hippuric acid, additionally identified by ANOVA as metabolites of interaction. 3,4-Dihydroxycinnamic acid (caffic acid) (Gulcin, 2006), and 2,4-Dihydroxybezoic acid (Sanz *et al*., 2012; Gorinstein *et al*., 1993), are also well known polyphenols found in wine. Hippuric acid on the other hand, is a normal human metabolite and is produced primarily in the liver by combination of dietary benzoic acid and glycine (Toromanovic *et al*., 2008). Van Dorsten *et al*, (2010), however reported an increase of 35% of the latter after wine consumption. Considering this evidence, it does seem as if the treatment failure group might have continuously consumed large amounts of wine during the treatment regimen, which may have contributed to TB treatment failure (Jaiswal *et al*., 2003).

Finally, metabolites associated with bacterial overgrowth, including 3-(3-hydroxyphenyl)-3-hydroxypropapoinc acid (HPHPA), were previously found to be elevated in the urine of the treatment failure group at time of diagnosis, as described in section 5.2.1.2.3. HPHPA was again identified as an important marker between the two outcome groups during treatment. Cis-4-hydroxycyclohexanecarboxylic acid has also previously been reported as a metabolite associated with bacterial overgrowth in the gastrointestinal tract (Kronick *et al*., 1983). These two metabolites, were found to be reduced in the urine in both groups during treatment, and this decrease was ascribed to administration of the potent anti-TB drugs eliminating the harmful intestinal bacteria. Furthermore, ANOVA analysis in this section once again

confirmed HPHPA and cis-4-hydroxycyclohexanecarboxylic acid, as important interaction markers between the two outcome groups. It was also confirmed by the time dependent analyses, that the greatest deference in these two metabolites when comparing the two outcome groups, were during the treatment phase of this study, with a general decrease observed in the treatment failure group over time (Appendix 2ah and 2ai).

Chapter 6

Conclusion

6.1 General conclusion

In conclusion, I will summarise the major findings made in Chapter 5 in the context of the aims proposed in Chapter 3. Additionally, I will discuss final remarks and future research prospects, based on the outcomes of this metabolomics study.

Although TB is considered a curable disease, it is one of the deadliest infectious diseases of our time, with 1.4 million deaths reported annually. Currently, South Africa ranks third on the 2011 list of 22 high-burden TB countries in the world and it was estimated that each active-TB person could potentially infect 10–15 people annually (WHO, 2011). Unfortunately, non-adherence to anti-TB treatment is still a major issue, due to it resulting in a global increased prevalence of drug resistant TB (WHO, 2011(b)). The high prevalence of treatment failure after the prescribed six month treatment regimen, further contributes to the above mentioned (WHO, 2011). It is thus clear, that greater efforts are urgently needed to strengthen the overall TB control in order to lower these TB incidences.

Thus, a great need exists for the identification of newer, distinctive, more sensitive and specific biomarkers, using new research approaches in order to better characterise TB treatment response and outcomes. In order to better explain the underlying mechanisms associated with this occurrence of treatment failure and to provide a better understanding to why certain individuals don't respond with a positive treatment outcome.

Consequently, we conducted a metabolomics research study using GC×GC/TOFMS to analyse the organic acid profiles of urine samples collected from patients clinically diagnosed with TB. Organic acid profiles of patients with a successful (cured) and unsuccessful (failed) treatment outcome were compared at various time points over the six months treatment regimen, in order to identify potential metabolite markers which may be used to better characterise the underlining mechanisms associated with poor treatment outcomes (treatment failure), ultimately aiding in alternative treatment approaches or interventions for these cases.

The first aim of this metabolomics study was to evaluate the accuracy and efficiency of the methodology used, as well as to determine the capability and accuracy of the analyst to perform these methods. In order to evaluate the GCxGC-TOFMS analytical repeatability, one QC sample was extracted and injected repeatedly (6 times) onto the GC×GC-TOFMS. Similarly, the analyst's repeatability for performing the organic acid analyses was also determined, using the 10 QC samples, which were extracted and injected separately. CV values were subsequently calculated from the collected and processed data, as a measure of this. As reported in Chapter 5, section 5.1.1, all of the 10 compounds selected for this evaluation, which were representative of the various compound classes and detected at regular retention time intervals throughout the total chromatographic run, fell within the predetermined cut-off of 20% for both the GCxGC-TOFMS and extraction method / analyst evaluations. Furthermore, 95.59% of all the compounds detected from the 6 QC sample repeats used for GCxGC-TOFMS repeatability, fell below a 50% CV value, and 93,7% of all the compounds analysed for analyst repeatability had a CV < 50%, which is the recommended cut-off from data matrix generation for metabolomics investigations. These results indicated that both GCxGC-TOFMS and the extraction method/analyst, have the capacity for generating repeatable results in the context of this metabolomics study.

Subsequently, using the above metabolomics approach, in addition to a wide variety of univariate and multivariate statistical methods, two patient outcome groups were compared, a sample group cured from TB after 6 months of treatment vs a sample group where treatment failed after the 6 month period at various time points: 1) time of diagnosis, before any anti-TB treatment were administered (objective 2), 2) during the course of treatment, in order to determine any variance in these groups due to a varying response to the anti-TB drugs (objective 3), 3) outcome groups were compared over all five time points at which urine were collected during the six month treatment regimen, in order to determine if differences exist between the two groups over time (objective 4).

A clear natural differentiation between the cured and failed outcome groups were obtained at time of diagnosis, and a total of 39 metabolites markers were subsequently identified. This result is astounding from both a biological point of view, considering that we may potentially explain why certain individuals would or would not respond to treatment, due to differences in their underlying metabolome, possibly due to certain genetic traits, as well as from a diagnostic perspective, potentially identifying those individuals who will respond to the anti-

TB medication and those who won't, even before treatment begins, who would potentially be placed on an alternative treatment strategy.

These metabolites are classified according to their various origins, and include (1) those associated with the presence of *M. tuberculosis* bacteria, (2) those resulting from an altered host metabolism due to the TB infection, and (3) metabolites of various exogenous origins. The detailed interpretation of these metabolites as described in section 5.2.1.2, suggests that a possible underlying RCD or some sort of mitochondrial dysfunction may be present in the treatment failure group. We hypothesise that this may possibly result in a far greater severity to *M. tuberculosis* infection in this group, subsequently causing a reduced capacity for a successful treatment outcome, also considering the critical role of the mitochondria in the metabolism of anti-TB drugs.

Furthermore, it is well known that infection with *M. tuberculosis* results in a host macrophage response, leading to production of ROS, causing oxidative stress and eventually resulting in alteration in the host metabolome. In the treatment failure group, TB infection led to a more distinct production of ROS and severity of oxidative stress, promoting inflammation and DNA damage. This ultimately results in ETC inhibition, decreased mitochondrial function, mark reduction in ATP production and consequently TCA cycle inhibition. As a result of this, impairment of mitochondrial fatty acid β-oxidation may follow, resulting in the production of dicarboxylic acids by means of ω-oxidation. Additionally, a greater severity of the oxidative stress and consequential increase H_2O_2 production, may lead to peroxisomal dysfunction/ degradation, and eventual impairment of α-oxidation.

Metabolite markers associated with elevated levels of various neurotransmitters (including GABA, dopamine and epinephrine) and amino acid metabolism (such as tryptophan, phenylalanine, tyrosine and isoleucine), were found in the treatment failure group, which further support the notion of an underlying RCD. Metabolites related to a possible pronounced hepatotoxicity or liver damage, as well as metabolites implicating a tendency towards elevated blood glucose levels and diabetes mellitus, were also found in the treatment failure group at time of diagnosis. Additional markers of liver damage were also observed in this outcome group during treatment, inductive of side effects caused by anti-TB drugs.

As mentioned in section 5.2.1.2.3, a variety of metabolites associated with exogenous origins, were found to be elevated in the treatment failure group. A large number of these were identified as metabolites indicative of microbial overgrowth in the gastrointestinal tract of these individuals. TB is known to disturb the normal microbial ecosystem in infected patients, thus it is possible that patients in the treatment failure group have a greater sensitivity to the *M. tuberculosis* infection, possibly due to the underlying RCDs. Other metabolites related to polyphenolic compounds commonly found in wine, were also found to be elevated in this group, and ascribed to increased wine consumption and the fact that these patients live in rural areas located in the Western Cape, where it is known that large amounts of wine is consumed. This may also be the cause of the RCD's detected, or further contribute to a susceptibility to this, in the treatment failed group.

Furthermore, 20 metabolite markers were identified when comparing the two outcome groups during the treatment phase of this metabolomics investigation (Objective 3). A vast majority of these 20 metabolites were also identified as markers for time 0 (at time of diagnosis). Metabolites indicative of impaired TCA cycle, reduced production of ATP, increased levels of neurotransmitters, impaired α-oxidation, as well as peroxisomal dysfunction, were elevated in the treatment failure group during treatment. It is concluded, that the ROS generated by anti-TB drugs, results in comparatively more severe metabolic alterations in the treatment failure group, probably due to the underlying RCD. Furthermore, metabolites associated with anti-TB drug induced side effects, were also found to be comparatively elevated in this group of patients, indicative of more pronounced liver damage in the treatment failure group and accompanied by metabolites characteristic of a MADD metabolite profile, due to a deficient electron transport flavoprotein, confirming previous experiments done in rats. These side effects have also previously been considered a major contributor of poor treatment compliance, which eventually will result in treatment failure. Additionally, a variety of exogenous markers were also identified which suggest more severe side effects in the treatment failure group. Metabolites indicative of additional medication such as aspirin taken by this patient group during treatment of anti-TB drugs were found to be elevated, probably to reduce the pain and symptoms associated with the anti-TB drug induced side effects. Wine metabolites were yet again found to be elevated in this patient group during the treatment phase, indicating that wine consumption continued during treatment of anti-TB drugs. Alcohol consumption is also known as a contributing factor of poor treatment compliance. Interestingly, elevated levels of ascorbic acid were detected in the cured patient group during this period, suggesting that ascorbic acid may aid in treatment.

Finally, 35 metabolite markers were identified by time dependent statistical analysis and represented those metabolites best describing the variation between the treatment outcome groups over the whole period (from diagnosis, to week 26). This time dependent statistical analysis was performed in order to confirm previous findings as well as identify additional markers which may help us better characterise treatment outcome. Of the 35 metabolites identified, 20 were previously identified and discussed in the previous investigations, confirming these findings. Additionally a large number of additional metabolites, characteristic of a MADD metabolic profile and liver damage were detected and elevated in the treatment failed group, which further supported the previous conclusions of more pronounced drug induced side effects in this patient group. Additional metabolites associated with oxidative stress, impaired TCA cycle and elevated lipid peroxidation, were also found to be elevated in the treatment failed group, further supporting the possibility of an underlying RCD and additional alterations caused by administration of anti-TB drugs in this group.

Considering the above, we successfully applied a metabolomics approach for identifying metabolites which could ultimately aid in the prediction and monitoring of treatment outcomes. This additionally lead to a better understanding and or characterisation of the phenomenon known as treatment failure, as well as the underlying mechanisms related to this occurrence.

6.2 Future Recommendations

Future prospects would aim towards the construction and validation of a diagnostic model which may be used to predict treatment outcome at time of diagnosis before anti-TB treatment is given. Using the 39 metabolites identified as metabolite markers at time of diagnosis, a characteristic biosignature for predicting treatment outcome may be compiled on the basis of a patient's characteristic metabolite profile, using the approach described by Olivier and Loots (2012). Furthermore, as indicated by the results obtained, underlying RCDs may be present in the treatment failed group. Further investigations may be conducted, using a more specific approach for testing for theses RCDs and the associated oxidative stress, in order to confirm our metabolomics generated hypotheses, formulated in this study. A number of unidentified metabolites were also identified as markers, especially when comparing the groups during the treatment phase of this experiment, most likely of

anti-TB drug origin, and thus not in the commercially available and in-house libraries. Further efforts could be made to identifying these, which may provide novel insight into possible variation in anti-TB drug metabolism and response in these two outcome groups, contributing to a better understanding of treatment failure.

Finally, a further investigation of the effects of alcohol consumption and vitamin C, on TB drug induced side effects and efficacy, is an important future prospect which emanated from this study.

Chapter 7

References

AHMAD, S. & MAKADDAS, E. 2010. Recent advances in the diagnosis and treatment of multidrug-resistant tuberculosis. *Respiratory Medicine CME,* 3:51-61.

ANNON. 2008. Isoniazid. *Tuberculosis,* 88(2):112-116.

ANTOINE, D., FRENCH, C.E., JONES, J. & WATSON, J.M. 2007. Tuberculosis treatment outcome monitoring in England, Wales and Northern Ireland for cases reported in 2001. *Journal of Epidemiol Community Health*, 61:302-307.

ARBEX, M.A., DE CASTRO LIMA VARELLA, M., DE SIQUERIA, H.R. & DE MELLO, F.A.F. 2010. Antituberculosis drugs: Drug interactions, adverse effects, and use in special situations. part 1: First-line drugs. *Jornal Brasileiro de Pneumologia,* 36(5):626-640.

ARSLAN, S., ÖZDEMIR, L., DEMIREL, Y. & AKKURT, I. 2010. The validity of the diagnostic methods in predicting pulmonary tuberculosis. *African Journal of Microbiology research,* 4(8):613-617.

ATUNA, R.A., BAEZAA, J., DROBNIEWSKIB, F., LEVICHEVAC, V. & COKERD, R.J. 2005. Implementing WHO DOTS strategy in the Russian federation: Stakeholder attitudes *Health policy,* 74:122-132.

BALAKRISHNAN, S., VIJAYAN, S., NAIR, S., SUBRAMONIAPILLAI, J., MRITHYNJAYAN, S., WILSON, N., SATYANARAYONA, S., DEWAN, P.K., KUMAR, A.M.V., KARTHICKEYAN, D., WILLIS, M., HARRIES, A.D. & NAIR, S.A. 2012. High diabetes prevalence among tuberculosis cases in Kerala, India. *PloS One,* 7(10):e46502.

BAKHSI, S. 2006. Tuberculosis in the United Kingdom: a tale of two nations. Trowbridge: Cromwell Press Ltd. 208 p.

BASSO, L.A. & SANTOS, D.S. 2005. Drugs that inhibit mycolic acid biosynthesis in *mycobacterium tuberculosis*-an update. *Medicinal Chemistry Reviews,* 2:393-413.

BELLO, S.I. 2010. Challenges of DOTS implementation strategy in the treatment of tuberculosis in a tertiary health institution, ilorin, nigeria. *African Journal of Pharmacy and Pharmacology,* 4(4):158-164.

BERNABE-ORTIZ, A., CARCAMO, C.P., SANCHEZ, J.F. & RIOS, J. 2011. Weight variation over time and its association with tuberculosis treatment outcome: a longitudinal analysis. *PLoS One*, 6(4):e18474.

BROUSSY, S., BERNARDES-GENISSON, V., GORNITZKA, H., BERNADOU, J. & MEUNIER, B. 2005. Studies on the 4-benzolpyridine-3-carboxamide entity as a fragment model of the isoniazid-NAD adduct. *Organic & Biomolecular Chemistry,* 3:666-669.

BULTITUDE, F.W. & NEWHAM, S.J. 1975. Identification of some abnormal metabolites in plasma from uremic subjects. *Clinical Chemistry,* 2:11329-1334.

CAMINERO, J.A., SOTGIU, G., ZUMLA, A. & MIGLIORI, G.B. 2010. Best drug treatment for multidrug-resistant and extensively drug-resistant tuberculosis. *Lancet Infectious Diseases,* 10:621-629.

CHUNG, H., YOON, Y.H., HWANG, J.J., CHO, K.S., KOH, J.Y. & KIM, J.G. 2008. Ethambutol-induced toxicity is mediated by zinc and lysosomal membrane permeabilization in cultured retinal cells. *Toxicology and Applied Pharmacology,* 235:163-170.

CRUZ, H.L.A., DE ALBUQUERQUE MONTENEGRO, R., DE ARAÚJO LIMA, J.F., DA ROCHA POROCA, D., DA COSTA LIMA, J.F., MONTENEGRO, L.M.L., CROVELLA, S. & SCHINDLER, H. 2011. Evaluation of a NESTED-PCR for *Mycobacterium tuberculosis* detection in blood and urine samples. *Brazilian Journal of Microbiology,* 42:321-329.

DELAINE, T., BERNARDES-GENISSON, V., QUEMARD, A., CONSTANT, P., MEUNIER, B. & BERNADOU, J. 2010. Development of isoniazid-NAD truncated adducts embedding a lipophilic fragment as potential bi-substrate InhA inhibitors and antimycobacterials. *European Journal of Medicinal Chemistry,* 45:4554-4561.

DROBNIEWSKI, F.A., CAWS, M., GIBSON, A. & YOUNG, D. 2003. Modern laboratory diagnosis of tuberculosis. *The Lancet Infectious Diseases,* 3:141-147.

DU PREEZ, I. & LOOTS, D.T. 2012. Altered fatty acid metabolism due to rafampicin- resistance conferring mutations in the rpoB gene of *Mycobacterium tuberculosis* : mapping the potential of pharmaco-metabolomics for global health and personalized medicine. *OMICS,* 16(11): 596-603.

DU PREEZ, I. & LOOTS, D.T. 2013. Detection limit for differentiating between various *Mycobacterium* species and *Pseudomonas aeruginosa* using gas chromatography-mass spectrometry (GC-MS) metabolomics: A comparison of two extraction methods. *African Journal of Microbiology*, 7(9): 797-801.

DU PREEZ, I. & LOOTS, D.T. 2013. New sputum metabolite markers implicating adaptations of the host to *Mycobacterium tuberculosis*, and vice versa. *Tuberculosis*, 93(3): 330-337.

DU, Z.X., ZHANG, H.Y., MENG, X., GUAN, Y. & WANG, H.Q. 2009. Role of oxidative stress and intracellular glutathione in the sensitivity to apoptosis induced by proteasome inhibitor in thyroid cancer cells. *Biomedical Central Cancer,* 9(56):1-11.

DUNN, W.B., BAILEY, N.J.C. & JOHNSON, H.E. 2005. Measuring the metabolome: Current analytical technologies. *Analyst,* 130:606-625.

FATIMA, N. 2009. Newer diagnostic techniques for tuberculosis. *Respiratory Medicine CME,* 2:151-154.

FAURHOLT-JEPSEN, D., RANGE, N., PRAYGOD, G., JEREMIAH, K., FAURHOLT-JEPSEN, M., AABYE, M.G., CHANGALUCHA, J., CHRISTENSEN, D.L., PIPPER, C.B., KRARUP, H., WITTE, D.R., ANDERSEN, A.B. & FRIIS, H. 2011. Diabetes is a risk factor for pulmonary tuberculosis: a case-control study from Mwanza, Tanzania. *PloS One,* 6(8):e24215.

FENG-ZENG, Z., LEVY, M.H. & SUMIN, W. 1997. Sputum microscopy results at two and three months predict outcome of tuberculosis treatment. *International Journal of Tuberculosis and Lung Disease*, 1(6):570-572.

FLYNN, J.L. 2004. Immunology of tuberculosis and implications in vaccine development. *Tuberculosis,* 84:93-101.

FRANKEL, E.N. Lipid Oxidation. 1980. *Progress in Lipid Research*. 19:1-22.

FRIEDEN, T.R., STERLING, T.R., MUNSIFF, S.S., WATT, C.J. & DYE, C. 2003. Tuberculosis. *The Lancet,* 362:887-899.

GONZALEZ-MARTIN, J., GARCIA-GARCIA, J.M., ANIBARRO, L., VIDAL, R., ESTEBAN, J., BLANQUER, R., MORENO, S. & RUIZ-MANZNO, J. 2010. Consensus document on the diagnosis, treatment and prevention of tuberculosis. *Archivos de Bronconeumologia,* 46(5):255-274.

GOPINATH, K. & SINGH, S. 2009. Urine as an adjunct specimen for the diagnosis of active pulmonary tuberculosis. *International Journal of Infectious Diseases,* 13:374-379.

GRAHAM, S.P. 2011. Treatment of paediatric TB: Revised WHO guidelines. *Paediatric Respiratory Reviews,* 12:22-26.

GRAY, J.W. 2004. Childhood tuberculosis and its early diagnosis. *Clinical Biochemistry,* 37:450-455.

GREEN, C., HUGGETT, J.F., TALBOT, M., P., REITHER, K. & ZUMLA, A.I. 2009. Rapid diagnosis of tuberculosis through the detection of mycobacterial DNA in urine by nucleic acid amplification methods. *Lancet Infectious Diseases,* 9:505-511.

GULBAY, B.E., GURKAN, O.U., YILDIZ, O.A., ONEN, Z.P., ERKEKOL, F.O., BACCIOGLU, A. & ACICAN, T. 2006. Side effects due to primary anti-tuberculosis drugs during the initial phase of therapy in 1149 hospitalized patients for tuberculosis. *Respiratory Medicine,* 100:1834-1842.

HESSELING, A.C., WALZL, G., ENARSON, D.A., CARROLL, N.M., DUCAN, K., LUKEY, P.T., LOMBARD, C., DONALD, P.R., LAWRENCE, K.A., GIE, R.P., VAN HELDEN, P.D. & BEYERS, N. 2010. Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *International Journal of Tuberculosis and Lung Disease,* 14(5):560-570.

HEWER, R., VOSTER, J., STEFFENS, F.E. & MEYER, D. 2006. Applying biofluid ^{1}H NMR-based metabonomic techniques to distinguish between HIV-1 positive/AIDS patients on antiretroviral treatment and HIV-1 negative individuals. *Journal of Pharmaceutical and Biomedical Analysis,* 41:1442-1446.

HORNE, D.J., ROYCE, S.E., GOOZE, L., NARITA, M., HOPEWELL, P.C., NAHID, P. & STEINGART, K.R. 2010. Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *Lancet Infectious Diseases*, 10:387-394.

HU, Y. & COATES, A.R.M. 2009. Acute and presistent *mycobacterium tuberculosis* infection depend on the thiol peroxidase TPX. *PLoS ONE,* 4(4):1-10.

HUO, T., CIA, S., LU, X., SHA, Y., YU, M. & LI, F. 2009. Metabonomic study of biochemical changes in the serum of type 2 diabetes mellitus patients after the treatment of metformin hydrochloride. *Journal of Pharmaceutical and Biomedical Analysis,* 49:976-982.

JABER, A.A., AHMAD, S. & MOKADDAS, E. 2009. Minor contribution of mutations at *iniA* codon 501 and *embC-embA* intergenic region in ethambutol-resistant clinical *mycobacterium tuberculosis* isolates in kuwait. *Annals of Clinical Microbiology and Antimicrobials,* 8(2):1-9.

JAIN, A. & MONDAL, R. 2008. Extensivelydrug-resistant tuberculosis: Current challenges and threats. *FEMS Immunology & Medical Microbiology,* 531:45-150.

JANIN, Y.L. 2007. Antituberculosis drugs: Ten years of research. *Bioorganic & Medicinal Chemistry,* 15:2479-2513.

KADDURAH-DAOUK, R., KRISTAL, B.S. & WEINSHILBOUM, R.M. 2008. Metabolomics: Aglobal biochemical approach to drug response and disease. *Annual Review of Pharmacology and Toxicology,* 48:653-683.

KALINOV´A, B., JIROˇS, P., ZD´´AREK, J., WEN, X. & HOSKOVEC, M. 2006. GC×GC/TOF MS technique-A new tool in identification of insect pheromones: Analysis of the persimmon bark borer sex pheromone gland. *Talanta,* 69:542-547.

KANEHISA, M., GOTO, S., KAWASHIMA, S., OKUNO, Y. & HATTORI, M. 2004. The KEGG resource for deciphering the genome. *Nucleic Acids rRsearch,* 32:277-280.

KARAKOUSIS, P.C. 2009. Mechanisms of action and resistance of antimycrobacterial agents. *Infectious Diseases,* 10:271-291.

KENCHEL, N.A. 2009. Tuberculosis: Pathophysiology, clinical features, and diagnosis. *Clinical Journal of The American Association of Critical-Care Nurses,* 29(2):34-43.

KIK, S.V., FRANKEN, W.P.J., MENSEN, M., COBELENS, F.G.J., KAMPHORST, M., AREND, S.M., ERKENS, C., GEBHARD, A., BORGDORFF, M.W. & VERVER, S. 2010. Predictive value for progression to tuberculosis by IGRA and TST in immigrant contacts. *European Respiratory Journal*, 35(6):1346-1353

KINDT, R., MOREEL, K., DEFORCE, D., BOERJAN, W. & VAN BOCXLAER, J. 2009. Joint GC-MS and LC-MS platforms for comprehensive plant metabolomics: Repeatability and sample pre-treatment. *Journal of Chromatography*, 877:3572-3580

KLOPPER, M., WARREN, R.M., HAYES, C., PITTIUS, N.C.G., STREACHER, E.M., MULLER, B., SIRGEL, A., CHABULA-NIXIWENI, M., HOOSAIN, E., COETZEE, G., VAN HELDEN, P.D. & VICTOR, T.C. 2013. Emergence and spread of Extensively and Totally Drug-Resistant Tuberculosis, South Africa. *Emerging Infectious Disease,* 19(3):449-455.

KONSTANTINOS, A. 2010. Testing for tuberculosis. *Australian Prescriber,* 33(1):12-18.

KOTLOWSKA, A., SWORCZAK, K. & STEPNOWSKI, P. 2011. Urine metabolomic analysis for adrenal incidentaloma activity detection and biomarker discovery. *Journal of Chromatography B,* 879:359-363.

KRIKHAM, P. 2007. Oxidative stress and macrophage function: A failure to resolve the inflammatory response. *Biochemical Society Transactions,* 35(2):284-287.

KUMAR, R. 2009. Glyoxylate shunt: Combating mycobacterium at fore front. *International Journal of Integrative Biology,* 7(2):69-72.

LAURENZO, D. & MOUSA, S.A. 2011. Mechanisms of drug resistance in *mycobacterium tuberculosis* and current status of rapid molecular diagnostic testing. *Acta Tropica,* 119(1):5-10.

LAWN, S.D., MWABA, P., BATES, M., PIATEK A., ALEXABDER, H., MARAIS, B.J., CUEVAS, L.E., McHUGH, T.D., ZIJENAH, L., KEPATA N., ABUBAKAR, I., MCNERNEY, R., HOELSCHER, M., MEMISH, Z.A., MIGLIORI, G.B., KIM, P., MAEURER, M., SCHITO, M. & ZUMLA, A. 2013. Advances in tuberculosis diagnostics: the Xpert MTB/RIF assay and future prospects for a point-of-care test. *Lancet Infectious Diseases,* 13(4):349-361.

LI, N., DENG, C. & ZHANG, X. 2007. Determination of methylmalonic acid and glutaric acid in urine by aqueous-phase derivarization followed by headspace solid-phase microextraction and gas chromatography-mass spectrometry. *Journal of Separation Science*. 30:266-271.

LOOTS, D.T., WIID, I.J., PAGE, B.J., MIENIE, L.J. & VAN HEERDEN, P.D. 2004. Melatonin prevents the free radical and MADD metabolic profiles induced by antituberculosis drugs in an animal model. *Journal of Pineal Research,* 38(2):100-106.

MADHUKAR PAI, M., KALANTRI, S., PASCOPELLA, L., RILEY, L.W. & REINGOLD, A.L. 2005. Bacteriophage-based assays for the rapid detection of rifampicin resistance in *mycobacterium tuberculosis*: A meta-analysis. *Journal of Infection,* 51:175-187.

MADSEN, R., LUNDSTEDT, T. & TRYGG, J. 2010. Cheometric in metabolomics-A review in human disease diagnosis. *Analytica Chimica Acta,* 659:23-33.

MAZUREK, G.H., LOBUE, P.A., DELAEY, C.L., BERNARDO, J., LARDIZABAL, A.A., BISHAI, W.R. IADEMARCO, M.F. & ROTHER, J.S. 2001. Comparison of a whole-blood interferon γ assay with tuberculin skin testing for detecting latent Mycobacterium tuberculosis infection. *JAMA*, 286(14):1740-1747.

MCCANCE-KATZ, E.F., MOODY, D.E., PRATHIKANTI, S., FRIEDLAN, G. & RAINEY, P.M. 2011. Rifampin, but not rifabutin, may produce opiate withdrawal in buprenorphine-maintained patients. *Drug and Alcohol Dependence,* 118(2-3):326-334.

MENZIES, D., BENEDETTI, A., PAYDAR, A., MARTIN, I., ROYCE, S., PAI, M., VERNON, A., LIENHARDT, C. & BURMAN, W. 2009. Effect of Duration and Intermittency of Rifampin on Tuberculosis Treatment Outcomes: A Systematic Review and Meta-Analysis. *PLoS Medicine*, 6(9): e1000146.

MORITA, K., ARIMOCHI, H. & OHNISHI, Y. 2003. In vitro cytotoxicity of 4-methylcatechol in murine tumor cells: Induction of apoptotic cell death by extracellular pro-oxidant action. *The Journal of Pharmacology and Experimental Therapeutics,* 306(1):317-323.

NICOL, M.P., WORKMAN, L., ISAACS W., MUNRO, J., BLACK, F., ELEY B., BOEHME, CC., ZEMANAY, W. & ZAR, H.J. 2011 Accuracy of the Xpert MTB/RIF test for the diagnosis of pulmonary tuberculosis in children

admitted to hospital in Cape Town, South Africa: a descriptive study. *Lancet Infectious Diseases*, 11(11): 819-824.

NUEDA. M.J., CARBONELL, J., MEDINA, I., DOPAZO, J. & CONESA, A. 2010. Serial Expression Analysis: a web toll for the analysis of serial gene expression data. *Nucleic Acids Research*. 38:239-245.

OLIVIER, I. & LOOTS, D.T. 2011. An overview of tuberculosis treatment and diagnostics. what role could metabolomics play. *Journal of Cell and Tissue Research,* 11(1):2655-2671.

OLIVIER, I & LOOTS, D.T. A metabolomics approach to characterise and identify various *Mycobacterium* species. 2012. *Journal of Microbiological Methods*. 88:419-426.

ORESIC, M. 2009. Metabolomics, a novel tool for studies of nutrition, metabolism and lipid dysfunction. *Nutrition, Metabolism & Cardiovascular Diseases,* 19:816-824.

ORMEROD, L.P. 2004. Respiratory tuberculosis. *Medicine,* 140-145.

PAI, M., FLORES, L.L., PAI, N., HUBBARD, A., RILEY, L.W. & COLFORD, J.M. 2003. Diagnostic accuracy of nucleic acid amplification tests for tuberculosis meningitis: a systematic review and meta-analysis. *Lancet Infectious Diseases*, 3:633-643.

PAI, M., KALANTRI, S. & DHEDA, K. 2006. New tools and emerging technologies for the diagnosis of tuberculosis: Part I. Latent tuberculosis. *Molecular Diagnostics*, 6(3):413-422. (a)

PALOMINO, J.C. 2009. Nonconvential and new methods in the diagnosis of tuberculosis: Feasibility and applicability in the field. *European Respiratory Journal,* 26:339-350.

PALOMINO, J.C., LEAO, S.C. & RITACCO, V. 2007. Tuberculosis: from basic science to patient care. Brazil. 678 p.

PARIDA, S.K. & KAUFMANN, S.H.E. 2010. The quest for biomarkers in tuberculosis. *Drug Discovery Today,* 15(4):148-157.

PEH, K.K. & YUEN, K.H. 1995. Development and in vitro evaluation of a novel multiparticulate matrix controlled release formulation of theophylline. *Drug Development and Industrial Pharmacy,* 21(13):1545-1555.

PERRIN, F.M.R., LIPMAN, M.C.I., MCHUGH, T.D. & GILLESPIE, S.H. 2007. Biomarker of treatment response in clinical trials of novel antituberculosis agent. *Lancet Infectious Diseases,* 7:481-490.

PERSICO, A.M. & NAPOLIONI, V. Urinary p-cresol in autism spectrum disorder. 2012. *Neurotoxicology and Teratology*. 36: 82-90.

RALSTON-HOOPER, K., HOPF, A., OH, C., ZHANG, X., ADAMEC, J. & SEPU´LVEDA, M.S. 2008. Development of GCxGC/TOF-MS metabolomics for use in ecotoxicological studies with invertebrates. *Aquatic Toxicology,* 88:48-52.

RAMASWAMY, S. & MUSSER, J.M. 1998. Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tubercle and Lung Disease*, 78:3-29.

RAYN, D., ROBARDS, K., PRENZLER, P.D. & KANDALL, M. 2011. Recent and potential development in the analysis of urine: A review. *Analytica Chimica Acta,* 684:17-29.

REDDY, Y.N., MURTHY, S.V., KRISHNA, D.R. & PRABHAKAR, M.C. 2004. Role of free radicals and antioxidants in tuberculosis patients. *Indian Journal of Tuberculosis,* 51(4):213-218.

REINECKE, J.C., KOEKEMOER, G., VAN DER WESTHUIZEN, F.H., LOUW, R., LINDEQUE, J.Z., MIENIE, L.J. & SMUTS, I. 2011. Metabolomics of urinary organic acids in respiratory chain deficiencies in children. *Metabolomics,* 7(2):1-24.

ROCKE, D.M. & LORENZATO, S. 1995. a two-component model for measurement error in analytical chemistry. *Technometrics,* 37:176-184.

RUEDA, C.M., MARIN, N.D., GARCIA, L.F. & ROJAS, M. 2010. Characterization of CD4 and CD8 T cells producing IFN-γ in human latent and active tuberculosis. *Tuberculosis,* 90(5):346-357.

SCHOEMAN, J.C., DU PREEZ, I. & LOOTS, D. 2012. A comparison of four sputum pre-extraction preparation methods for identifying and characterising *Mycobacterium tuberculosis* using GCxGC-TOFMS metabolomics. *Journal of Microbiological Methods*, 91:301-311

SEKADDE, M.P., WOBUDEYA, E., JOBBA, M., SSENGOOBA, W., KRISEMBA, H., BAKEERA-KITAKA, S. & MUSOKE, P. 2013. Evaluation of the Xpert MTB/RIF test for the diagnosis of childhood pulmonary tuberculosis in Uganda: across-sectional diagnostic study. *BCM Infectious Diseases,* 13(133): 1-8

SHEEN, P., FERRER, P., GILMAN, R.H., PEZ-LLANO, J.L. & FUENTES, P. 2009. Effect of pyrazinamidase activity on pyrazinamide resistance in *mycobacterium tuberculosis*. *Tuberculosis,* 89:109-113.

SHI, R., ITAGAKI, N. & SUGAWARA, I. 2007. Ovweview of anti-tuberculosis (TB) drugs and their resistance mechanisms. *Mini-Reviews in Medicinal Chemistry,* 7:1177-1185.

SMITH, D.G., GUILLEMIN, G.J., PEMBERTON, L., KERR, S., NATH, A., SMYTHE, A.G. & BREW, B.J. 2001. Quinolinic acid is produced by macrophages stimulated by platelet activating factor, nef and tat. *Journal of Neurovirology,* 7:56-60.

TANAKA, T., SAKURADA, S., KANO, K., TAKAHASHI, E., YASUDA, K., HIRANO, H., KABURAGI, Y., KOBAYASHI, N., LE HANG, N.T., LIEN, L.T., MATSUSHITA, I., HIJIKATA, M., UCHIDA, T. & KEICHO, N. 2011. Identification of tuberculosis-associated proteins in whole blood supernatant. *Infectious Diseases,* 11(72):1-10.

THERON, G., PETER, J., VAN ZYL SMIT, R., MISHRA, H., STREICHER, E., MURRAY, S., DAWSON, R., WHITELAW, A., HOELSCHER, M., SHARMA, S., PAI, M., WARREN, R. & DHEDA, K. 2011. Evaluation of the Xpert MTB/RIF Assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. *American Journal of Respiratory and Critical Care Medicine,* 184:132-144.

TIAN, R.R., PAN, Q.H., ZHAN, J.C., LI, J.M., WAN, S.B., ZHANG, Q.H. & HUANG, W.D. 2009. Comparison of phenolic acids and flavan-3-ols during wine fermentation of grapes with different harvest times. *Molecules,* 14:827-838.

TOSTMANN, A., BOEREE, M.J., AARNOUTSE, R.E., DE LANGE, W.C.M., VAN DER VENS, A.J.A.M. & DEKHUIJZEN, R. 2007. Antituberculosis drug-induced hepatotoxicity: concise up-to-date review. *Journal of Gastroenterology and Hepatology,* 23:192-202.

TUFARIELLO, J.M., CHAN, J. & FLYN, J.L. 2003. Latent tuberculosis: Mechanisms of host and bacillus that contribute to persistent infection. *The Lancet Infectious Diseases,* 3:578-590.

VAN DORSTEN, F.A., GRUN, C.H., VAN VELZEN, E.J.J., JACOBS, D.M., JACOBS, D.M., DRAIJER, R. & VAN DUYNHOVEN, J.P.M. 2009. The metabolic fate of wine and grape juice polyphenols in humans assessed by metabolomics. *Molecular Nutrition & Food Research,* 54:897-908.

VELAYATI, A.A., FARNIA, P., MASJEDI, M.R., IBRAHIM, T.A., TABARSI, P., HAROUN, R.Z., KUAN, H.O., GHANAVI, J., FARNIA, P. & VARAHRAM, M. (2009). Totally drug-resistant tuberculosis strains: evidence of adaptation at the cellular level. *European Respiratory Journal*, 34(5): 1201-1203.

VILLANO, D., FERNANDEZ-PACHON, M.S., TRONCOSO, A.M. & GARCIA-PARRILLA, M.C. 2005. Comparison of antioxidant activity of wine phenolic compounds and metabolites in vitro. *Analytica chimica Acta,* 538:391-398.

VOLMINK, J. & GARNER, P. 1997. Systematic review of randomised controlled trials of strategies to promote adherence to tuberculosis treatment. *British Medical Journal*, 315:1403-1406.

WANCHU, A. 2005. Advances in serology for diagnosing TB in the HIV infected. *Indian Journal of Chest Diseases and Allied Sciences,* 47:31-37.

WEIS, S.E., SLOCUM, P.C., BLAIS, F.X., KING, B., NUNN, M., MATNEY, G.B., GOMEZ, E. & FORESMAN, B.H. 1994. The effect of directly observed therapy on the rates of drug resistance and relapse in tuberculosis. *The New England Journal of Medicine*; 330:1179-1184.

WILLEY, J.M., SHERWOOD, L.M. & WOOLVERTON, C.J. 2008. Prescott, Harley, and Klein's Microbiology. Singapore: McGraw-Hill. 1088 p.

WORLD HEALTH ORGANIZATION. 2011. Global tuberculosis control: WHO report 2011. Switzerland: World Health Organization. 258p. (a)

WORLD HEALTH ORGANIZATION. 2011. World Health Organization 2011/2012 tuberculosis global facts. Switzerland: World Health Organization. 2p. (b)

WORLD HEALTH ORGANIZATION. 2010. Roadmap for rolling out Xpert MTB/RIF for rapid diagnosis of TB and MDR-TB. Genena, Switzerland: World Health Organization.

WOLINSKY, E. 1994. Conventional diagnostic methods for tuberculosis. *Clinical Infectious Disease,* 19(3):396-401.

WU, H., XUE, R., DONG, L., LUI, T., DENG, C., ZENG, H. & SHEN, X. 2009. Metabolomic profiling of human urine in heptocellular carcinoma patients using gas chromatography/ mass spectrometry. *Analytica Chimica Acta,* 648:98-104.

WU, H., XUE, R., TANG, Z., DENG, C., LUI, T., ZENG, H., SUN, Y. & SHEN, X. 2010. Metabolomic investigation of gastric cancer tissue using gas chromatography/ mass spectrometry. *Analytical and Bioanalytical Chemistry,* 396:1385-1395.

XIA, J., PSYCHOGIOS, N., YOUNG, N. & WISHART, D.S. 2009. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acid Research,* 37:652-660.

XIA, J. & WISHART, D.S. 2011. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nature Protocols,* 6(6):743-760.

ZAUGG, S., ZHANG, X., SWEEDLER, J. & THORMANN, W. 2001. Determination of salicylate, gentisic acid and salicyluric acid in human urine by capillary electrophoresis with laser-induced fluorescence detection. *Journal of Chromatography B Biomedical Sciences and Applications,* 752(1):17-31.

ZHANG, Y. 2005. The magic bullets and tuberculosis drug targets. *Annual Review of Pharmacology and Toxicology,* 45:529-564.

Appendix 1

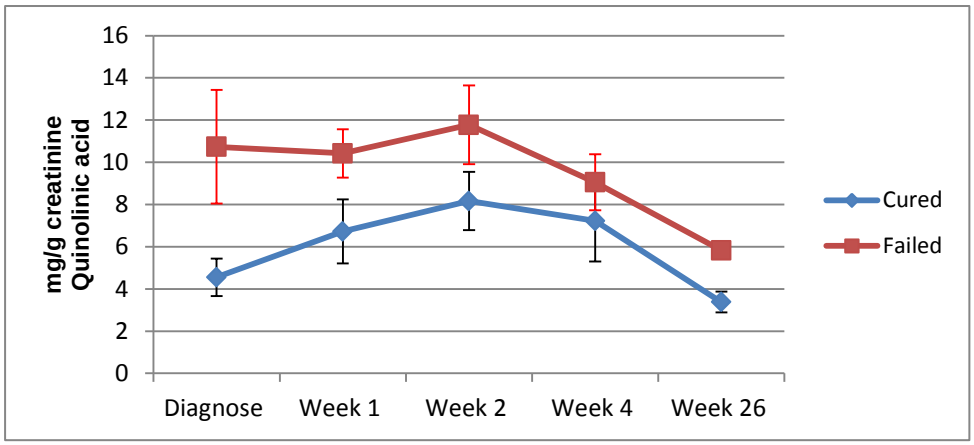
Table 1: Clinical and microbial information of the anonyms patients of which urine samples were collected and used in this metabolomic study.

Patient no.	Gender	Age	BMI	Bacteriological Culture		Sputum smears microscopy		Outcome
				Diagnosis	Week 26	Diagnosis	Week 26	
1	Male	58	17.689	Positive	Positive	0	0	Failed
2	Female	35	21.193	Positive	Positive	1	0	Failed
3	Male	28	19.541	Positive	Positive	0	0	Failed
4	Female	44	16.574	Positive	Positive	1	1	Failed
5	Male	42	15.02	Positive	Positive	1	0	Failed
6	Male	37	19.433	Positive	Positive	1	0	Failed
7	Male	53	15.967	Positive	Positive	1	0	Failed
8	Female	22	19.289	Positive	Positive	1	0	Failed
9	Female	36	18.084	Positive	Negative	1	0	Cured
10	Female	45	18.902	Positive	Negative	1	0	Cured
11	Female	25	18.612	Positive	Negative	1	0	Cured
12	Male	49	20.796	Positive	Negative	1	0	Cured
13	Male	39	16.298	Positive	Negative	0	0	Cured
14	Male	48	19.467	Positive	Negative	1	0	Cured
15	Male	42	15.887	Positive	Negative	1	0	Cured
16	Male	24	17.156	Positive	Positive	1	0	Failed
17	Female	28	21.471	Positive	Negative	0	0	Cured
18	Male	44	17.722	Positive	Negative	1	0	Cured
19	Male	37	15.78	Positive	Negative	2	0	Cured
20	Female	28	18.094	Positive	Negative	3	0	Cured
21	Male	25	16.184	Positive	Negative	1	0	Cured
22	Male	23	17.993	Positive	Negative	3	0	Cured
23	Male	31	15.571	Positive	Negative	3	0	Cured
24	Male	38	20.55	Positive	Negative	1	0	Cured
25	Female	17	22.136	Positive	Negative	1	0	Cured
26	Male	37	19.204	Positive	Positive	2	3	Failed
27	Female	55	20.078	Positive	Positive	1	1	Failed
28	Female	25	18.315	Positive	Positive	1	1	Failed
29	Male	40	20.45	Positive	Positive	1	0	Failed
30	Male	31	17.289	Positive	Positive	1	2	Failed
31	Female	45	18.591	Positive	Negative	1	0	Cured
32	Female	25	21.052	Positive	Negative	1	0	Cured
33	Male	24	18.819	Positive	Negative	1	0	Cured
34	Female	42	17.481	Positive	Negative	1	0	Cured
35	Female	37	18.518	Positive	Negative	2	0	Cured
36	Female	21	22.547	Positive	Negative	3	0	Cured
37	Female	25	18.905	Positive	Positive	1	1	Failed
38	Male	50	20.549	Positive	Negative	3	0	Cured
39	Female	54	16.359	Positive	Negative	0.3	0	Cured
40	Female	26	17.506	Positive	Negative	2	0	Cured

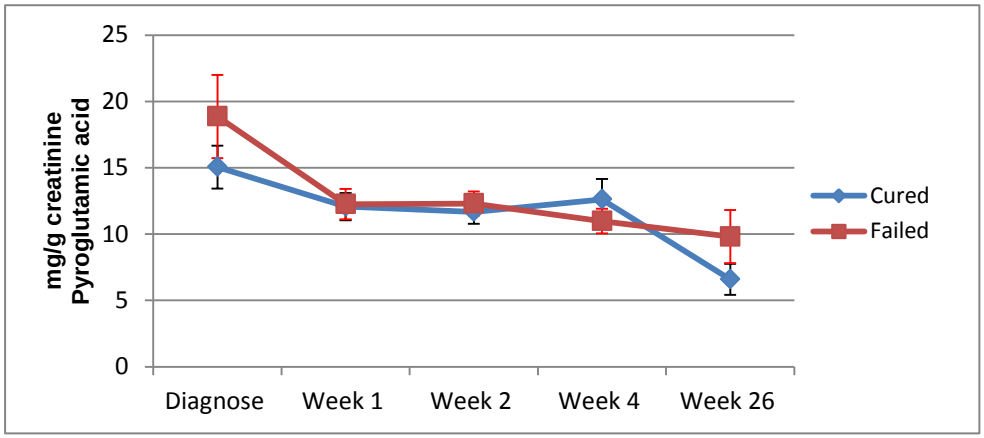
41	Female	45	21.967	Positive	Negative	0	0	Cured
42	Female	24	16.604	Positive	Negative	1	0	Cured
43	Male	31	16.975	Positive	Negative	1	0	Cured
44	Male	31	18.518	Positive	Negative	1	0	Cured
45	Male	21	19.265	Positive	Positive	2	1	Failed
46	Male	48	19.789	Positive	Positive	0	1	Failed

Appendix 2

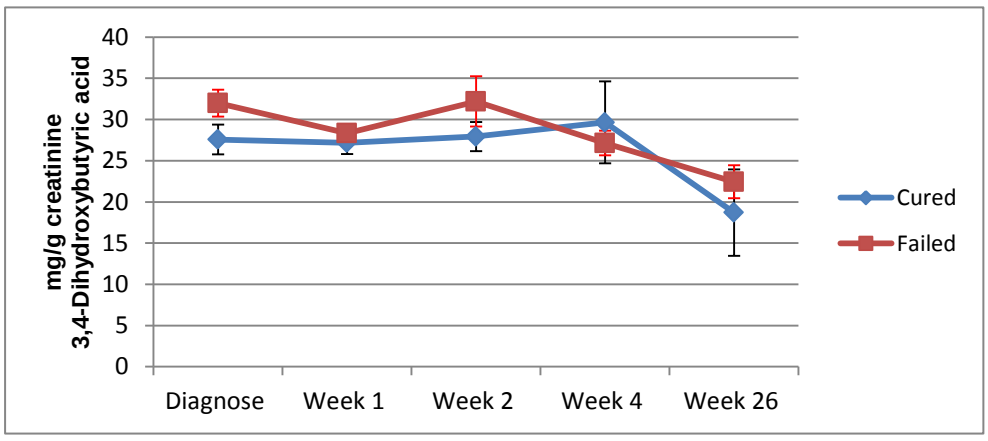
a


mg/g creatinine
Quinolinic acid
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

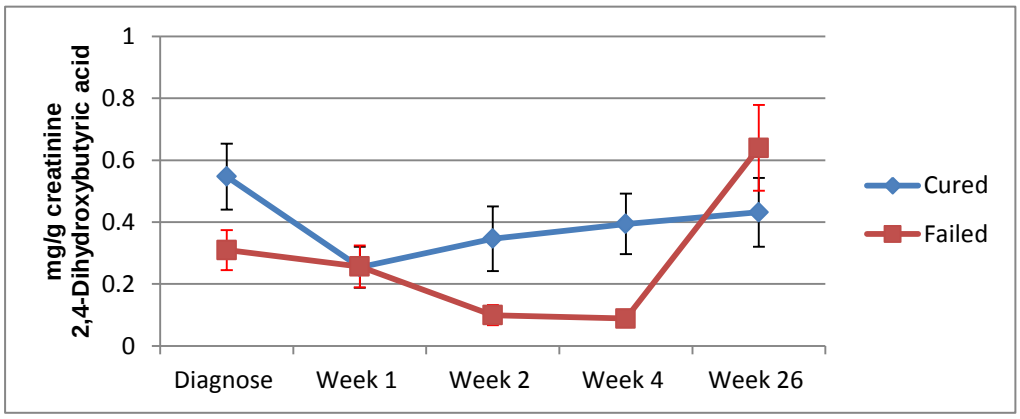
b


mg/g creatinine
Pyroglutamic acid
25
20
15
10
5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

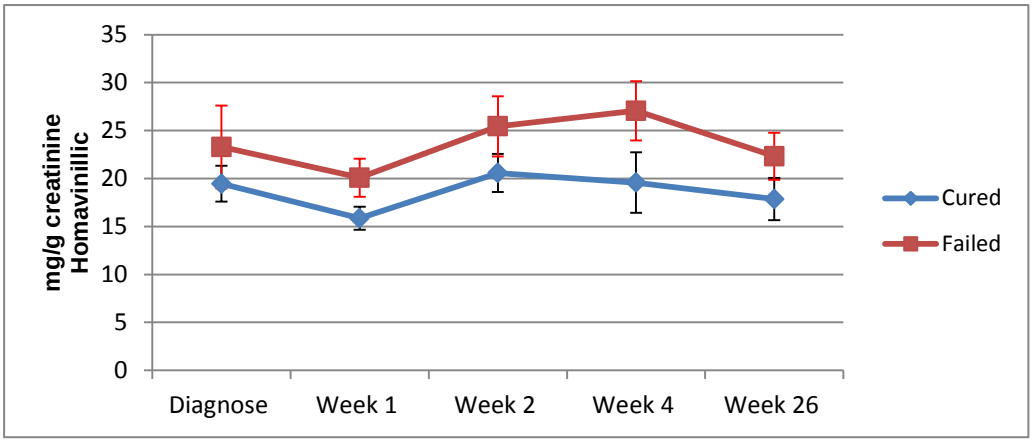
c


mg/g creatinine
3,4-Dihydroxybutyric acid
40
35
30
25
20
15
10
5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

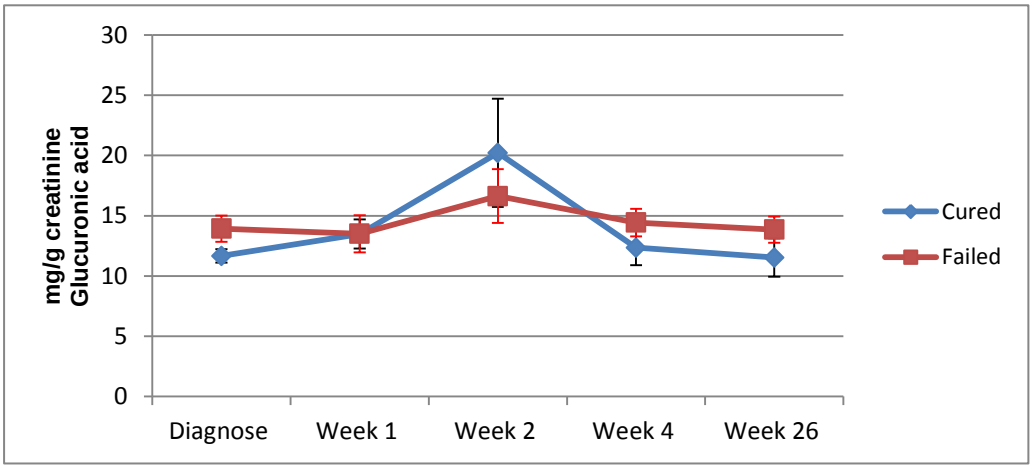
d


mg/g creatinine
2,4-Dihydroxybutyric acid
1
0.8
0.6
0.4
0.2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

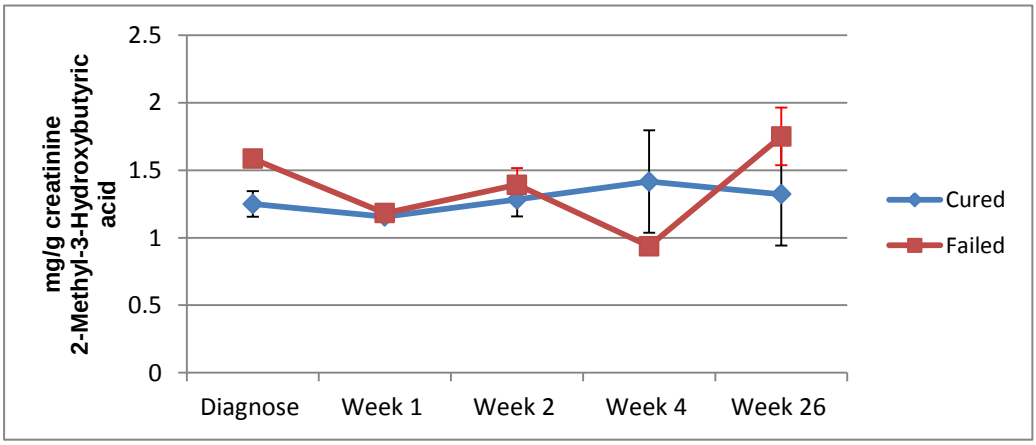
e


mg/g creatinine
Homavinillic
35
30
25
20
15
10
5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

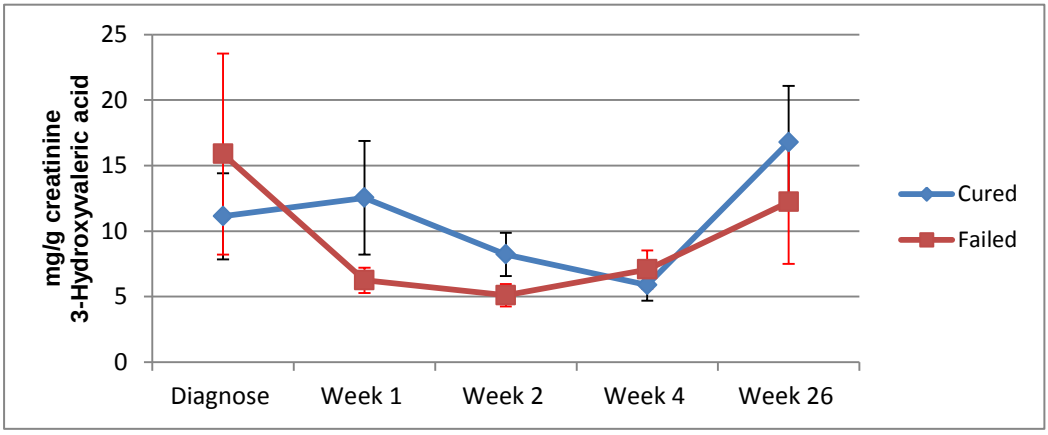
f


mg/g creatinine
Glucuronic acid
30
25
20
15
10
5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

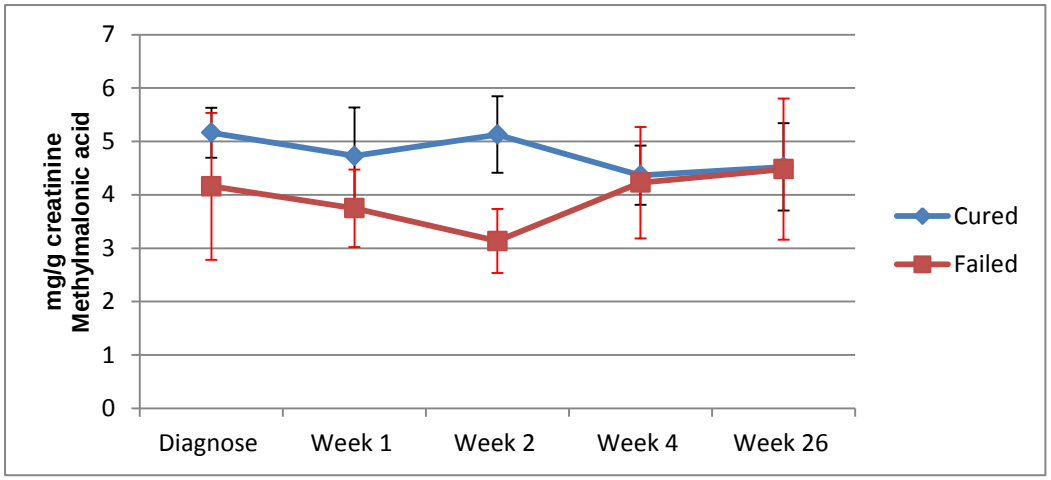
g


mg/g creatinine
2-Methyl-3-Hydroxybutyric acid
2.5
2
1.5
1
0.5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

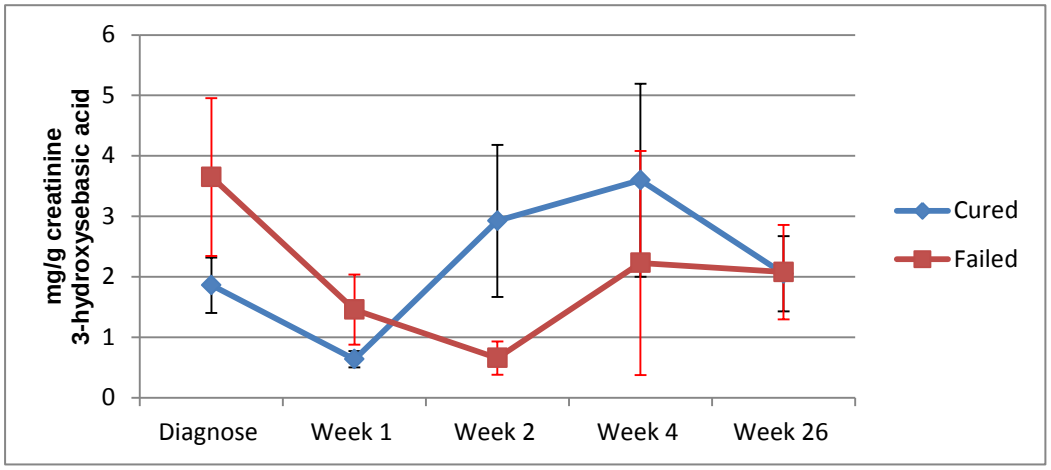
h


mg/g creatinine
3-Hydroxyvaleric acid
25
20
15
10
5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

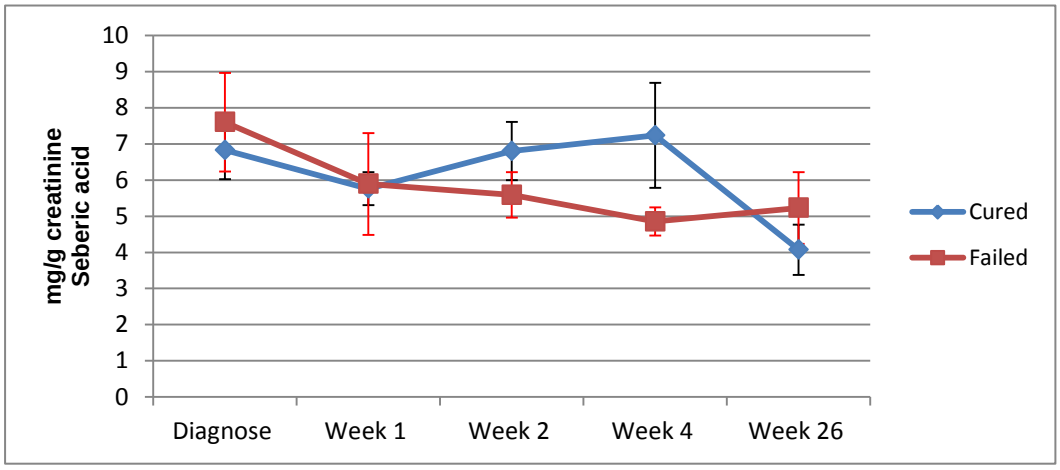
i


mg/g creatinine
Methylmalonic acid
7
6
5
4
3
2
1
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

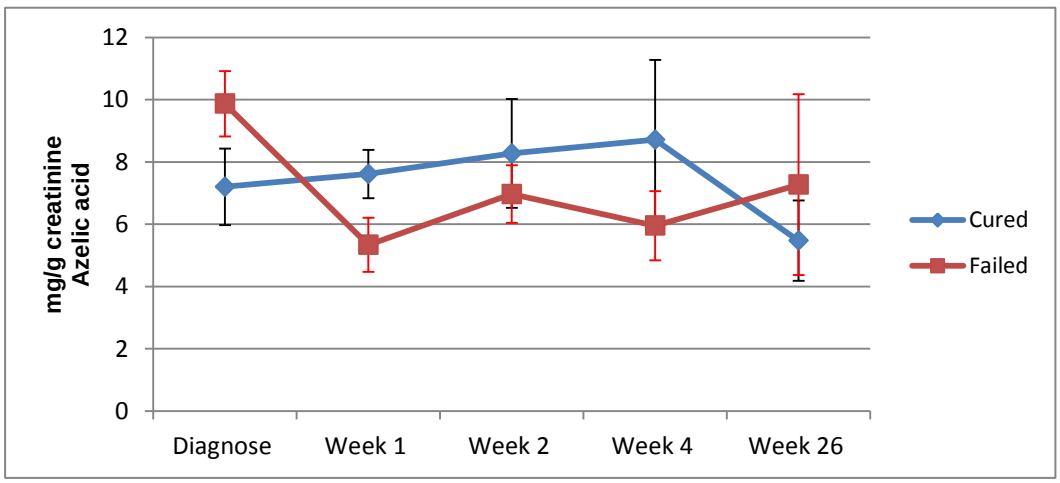
j


mg/g creatinine
3-hydroxysebasic acid
6
5
4
3
2
1
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

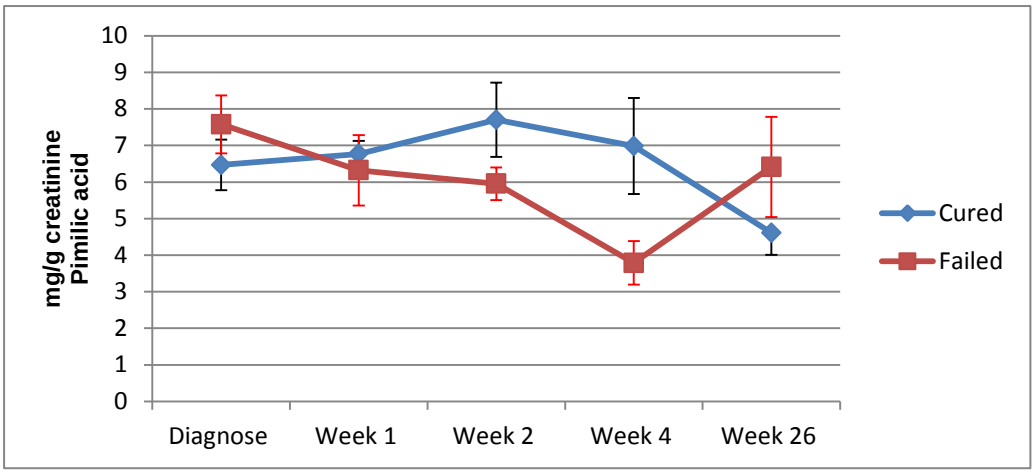
k


mg/g creatinine
Seberic acid
10
9
8
7
6
5
4
3
2
1
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

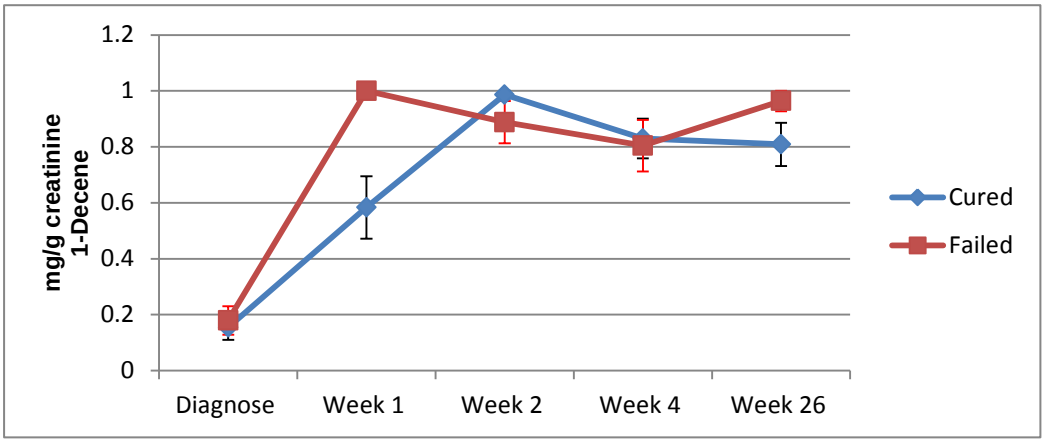
l


mg/g creatinine
Azelic acid
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

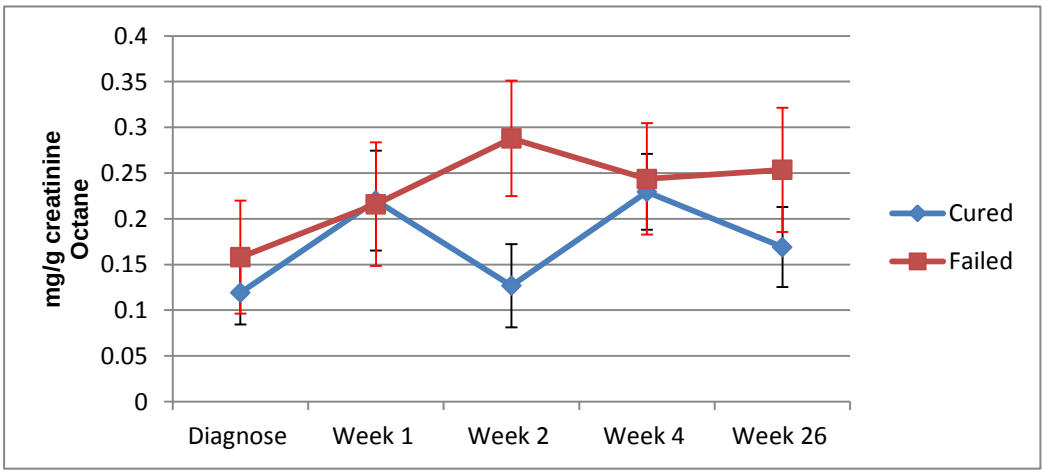
m


mg/g creatinine
Pimilic acid
10
9
8
7
6
5
4
3
2
1
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

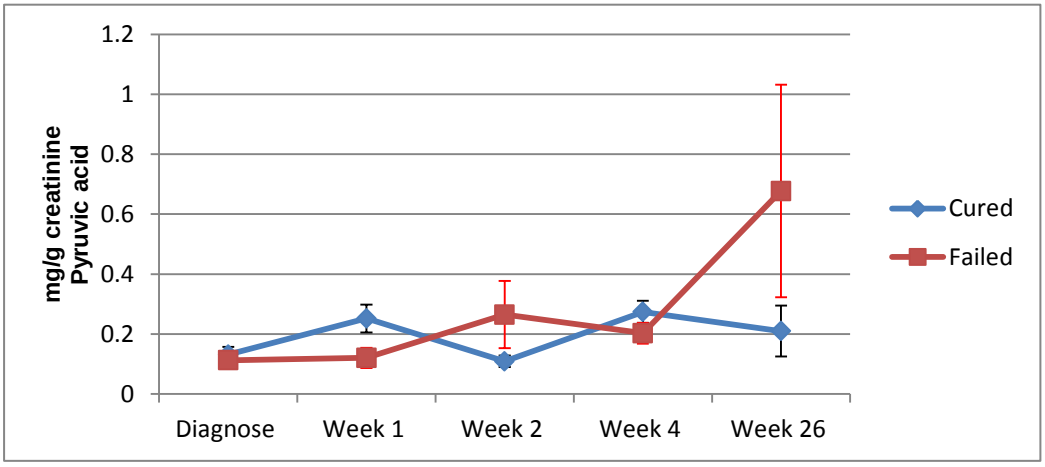
n


mg/g creatinine
1-Decene
1.2
1
0.8
0.6
0.4
0.2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

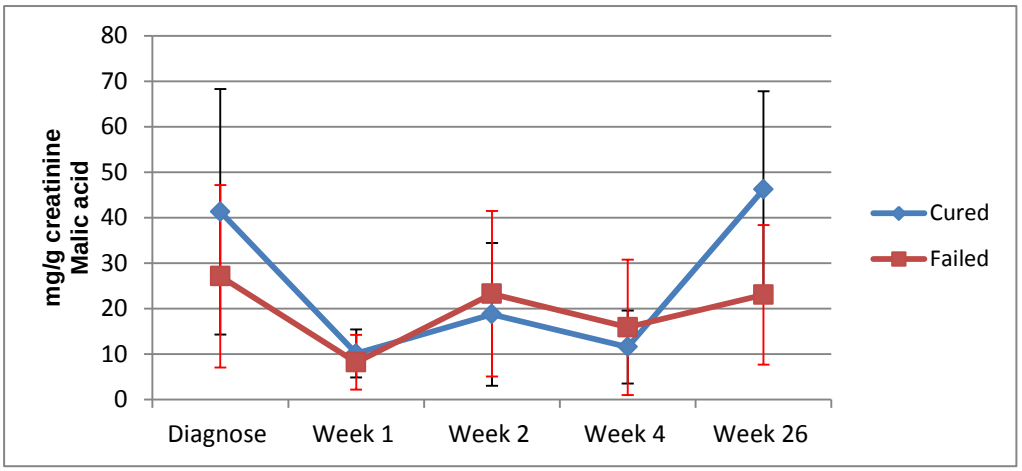
O


mg/g creatinine
Octane
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

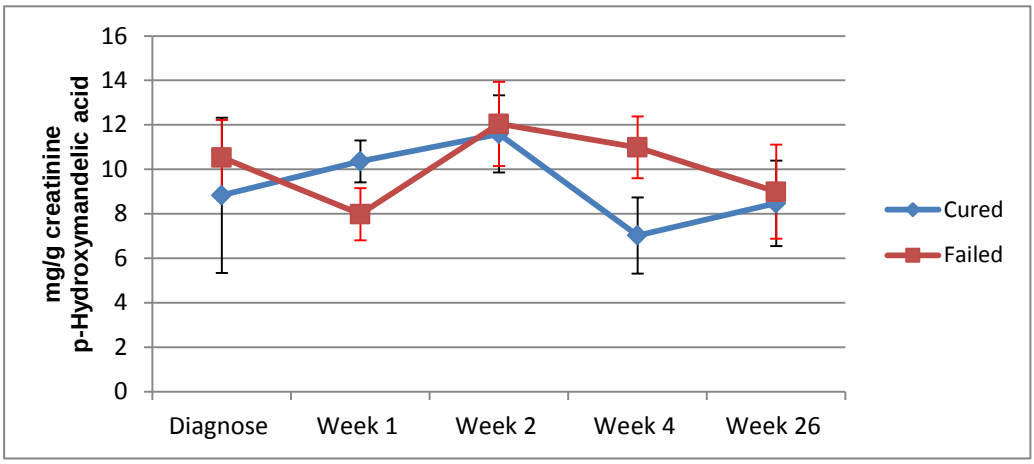
p


mg/g creatinine
Pyruvic acid
1.2
1
0.8
0.6
0.4
0.2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

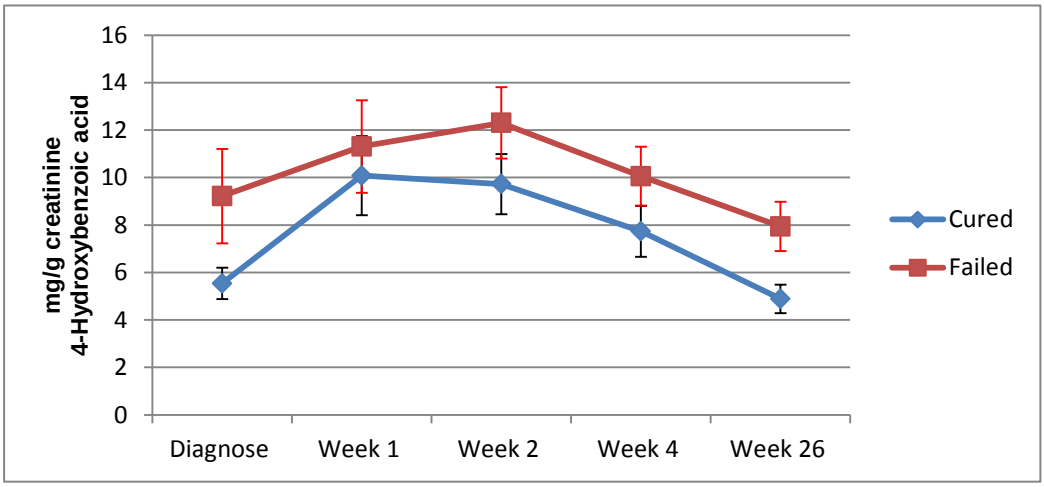
q


mg/g creatinine
Malic acid
80
70
60
50
40
30
20
10
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

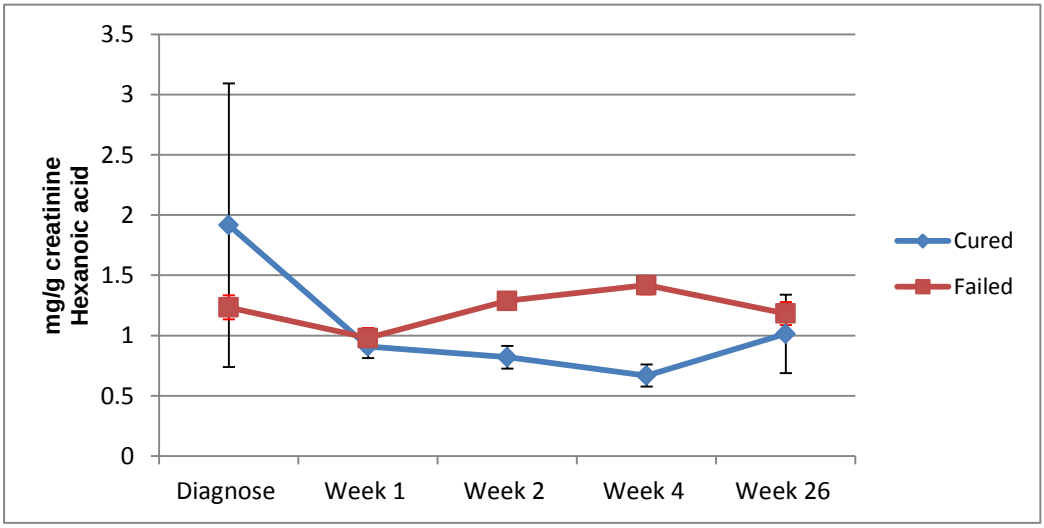
r


mg/g creatinine
p-Hydroxymandelic acid
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

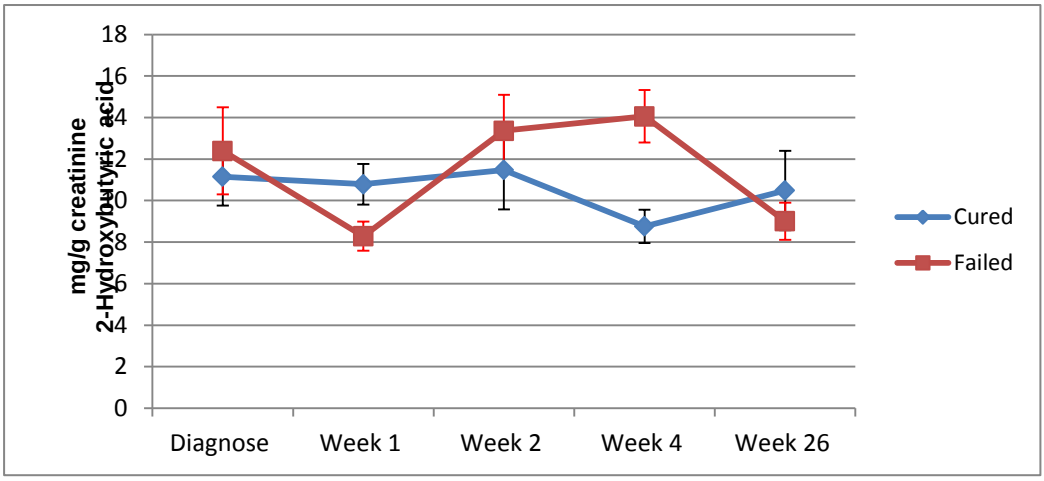
s


mg/g creatinine
4-Hydroxybenzoic acid
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

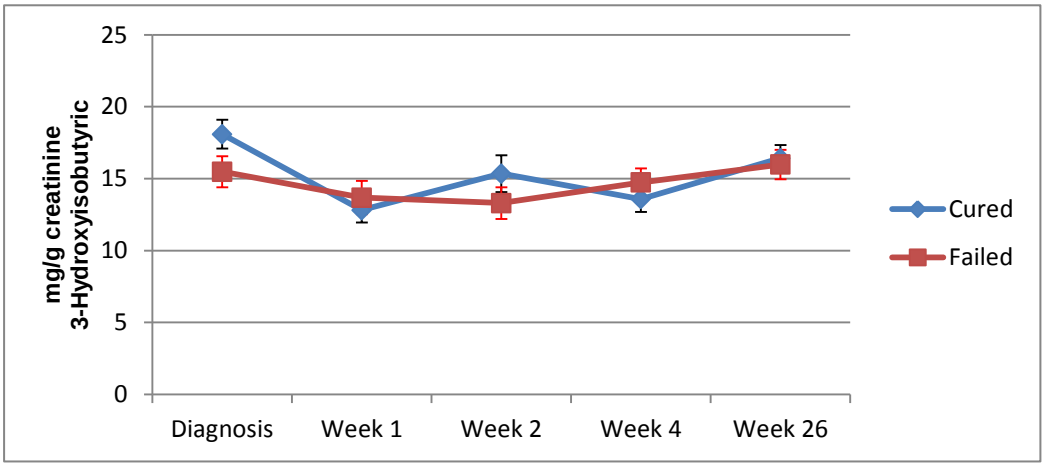
t


mg/g creatinine
Hexanoic acid
3.5
3
2.5
2
1.5
1
0.5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

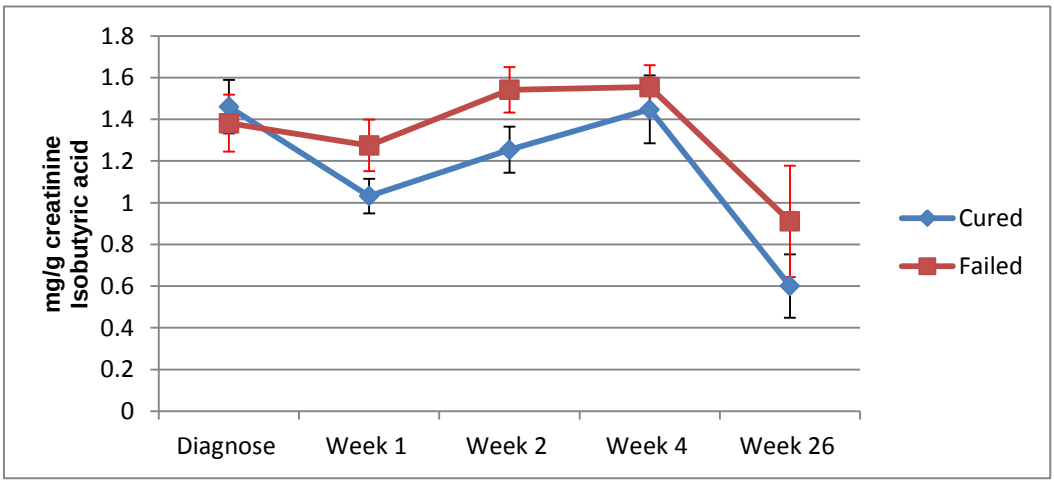
u


mg/g creatinine
2-Hydroxybutyric acid
18
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

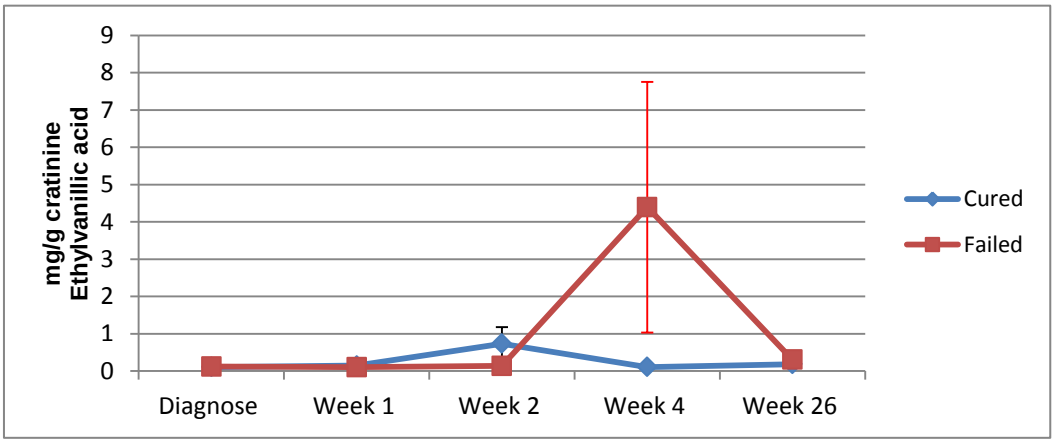
v


mg/g creatinine
3-Hydroxyisobutyric
25
20
15
10
5
0
Diagnosis
Week 1
Week 2
Week 4
Week 26
Cured
Failed

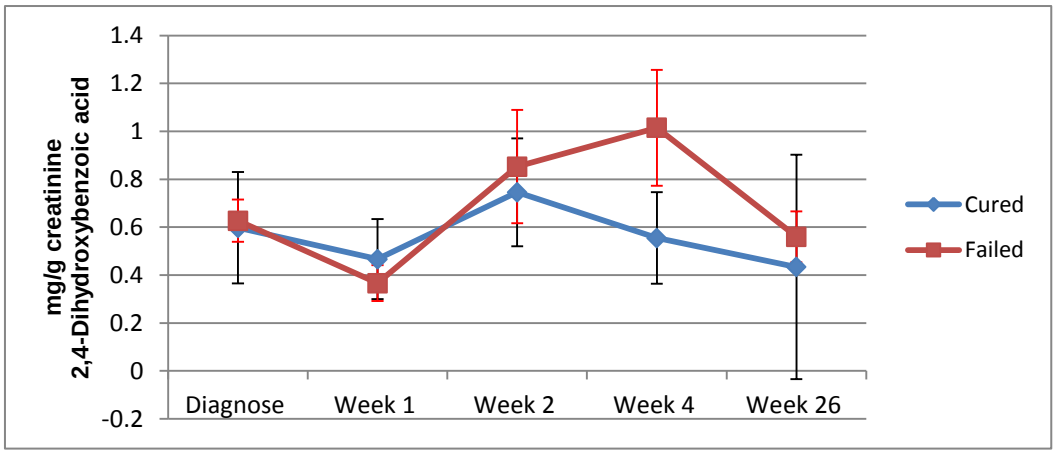
w


mg/g creatinine
Isobutyric acid
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

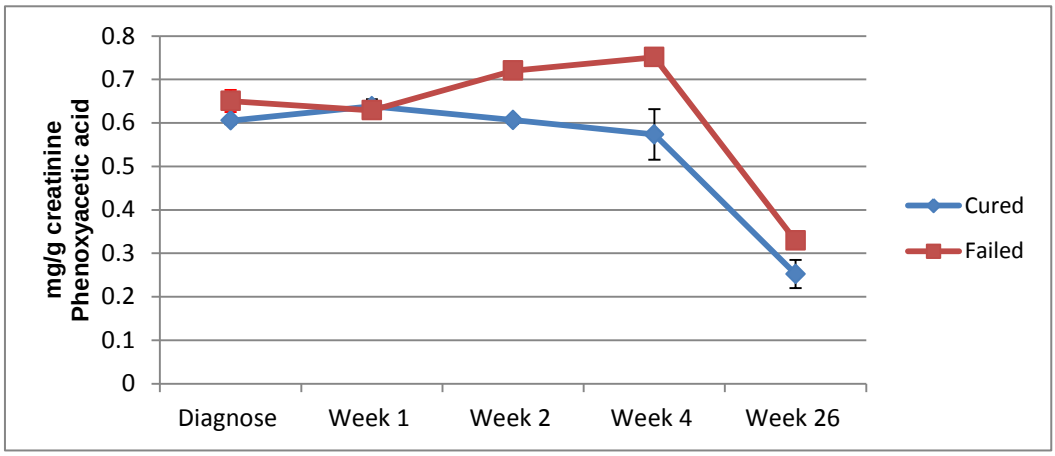
x


mg/g cratinine
Ethylvanillic acid
9
8
7
6
5
4
3
2
1
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

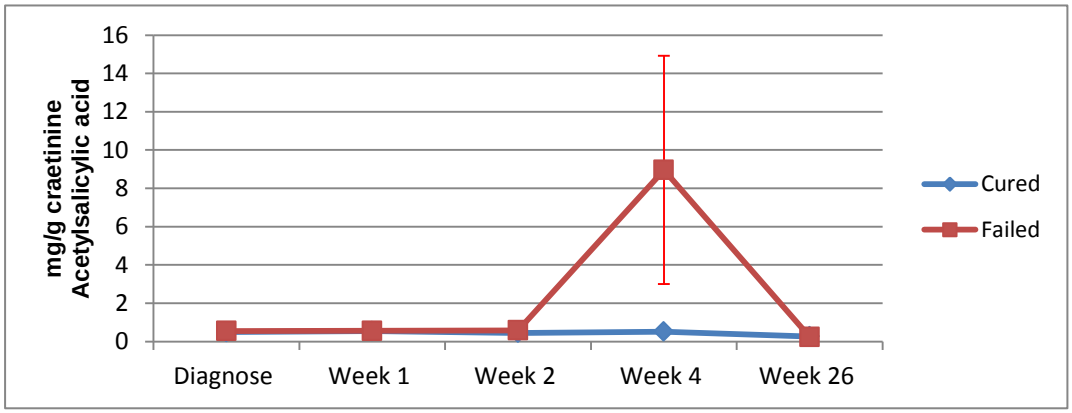
y


mg/g creatinine
2,4-Dihydroxybenzoic acid
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

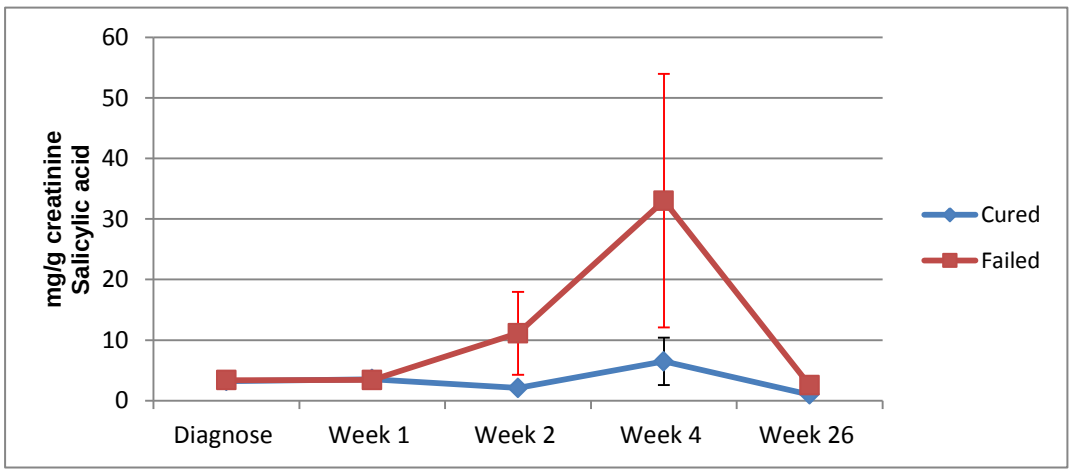
z


mg/g creatinine
Phenoxyacetic acid
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

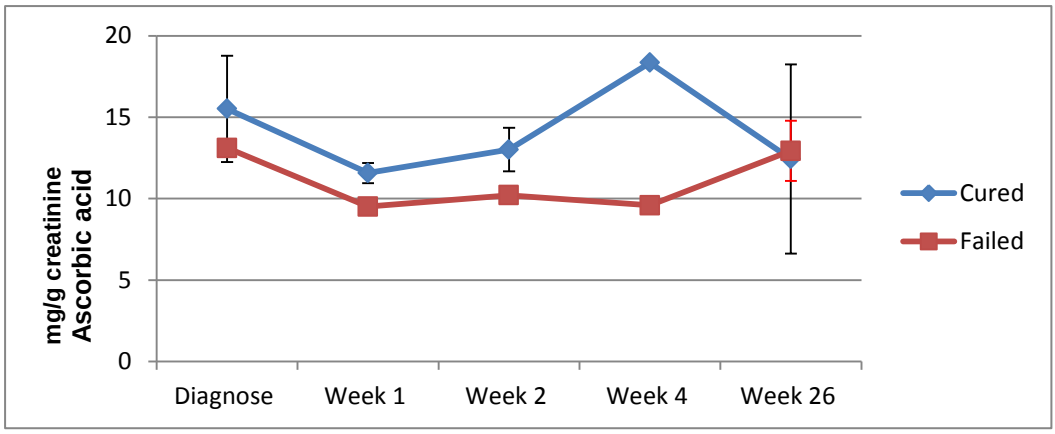
aa


mg/g craetinine
Acetylsalicylic acid
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

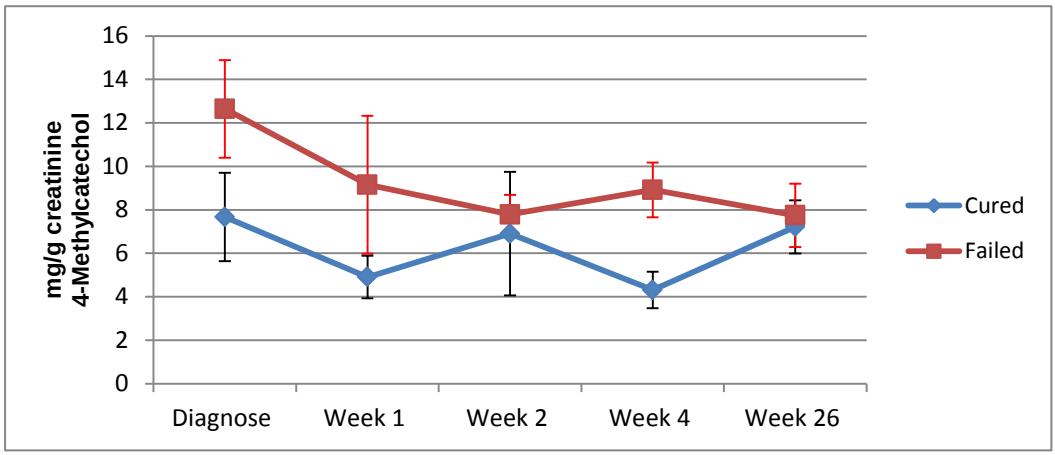
ab


mg/g creatinine
Salicylic acid
60
50
40
30
20
10
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

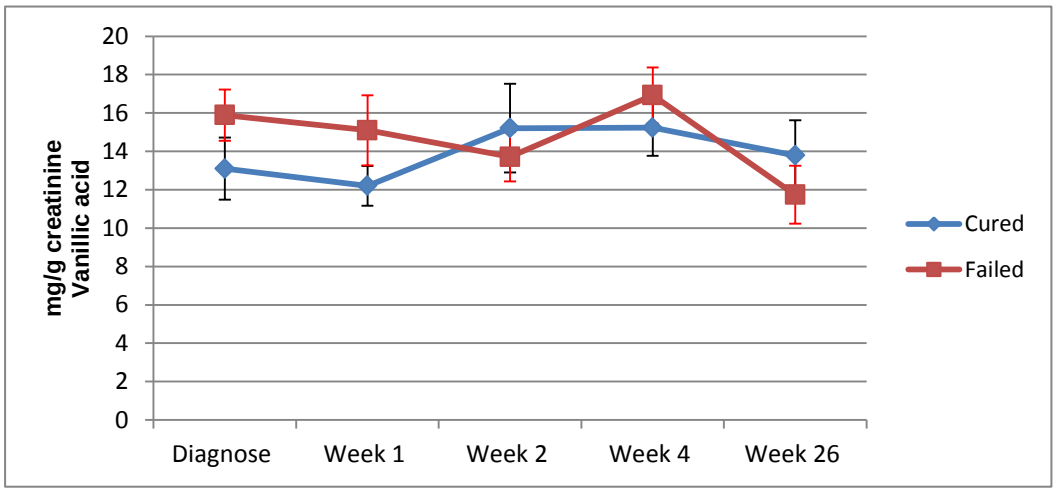
ac


mg/g creatinine
Ascorbic acid
20
15
10
5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

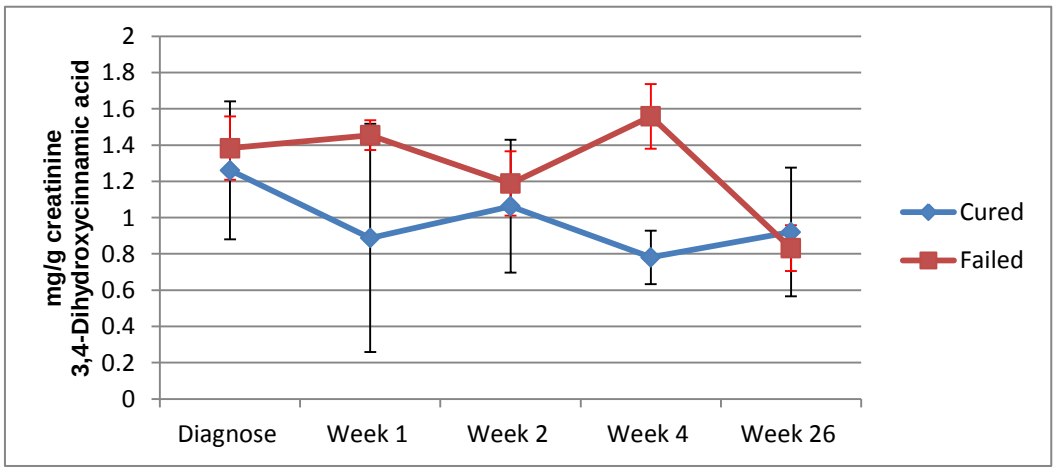
ad


mg/g creatinine
4-Methylcatechol
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

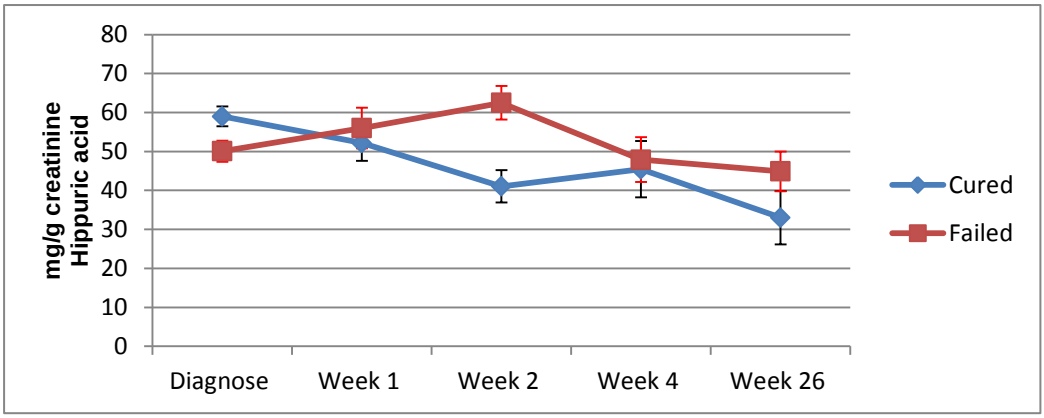
ae


mg/g creatinine
Vanillic acid
20
18
16
14
12
10
8
6
4
2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

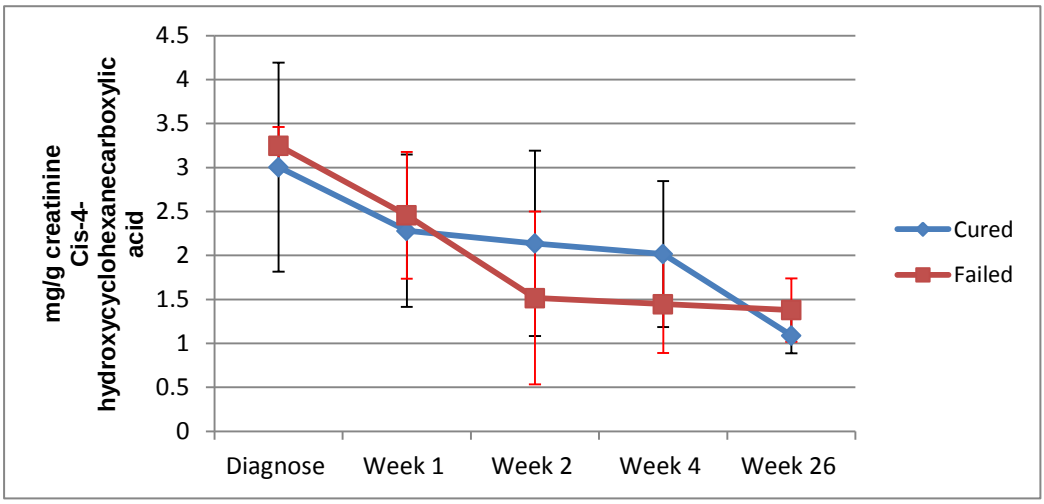
af


mg/g creatinine
3,4-Dihydroxycinnamic acid
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

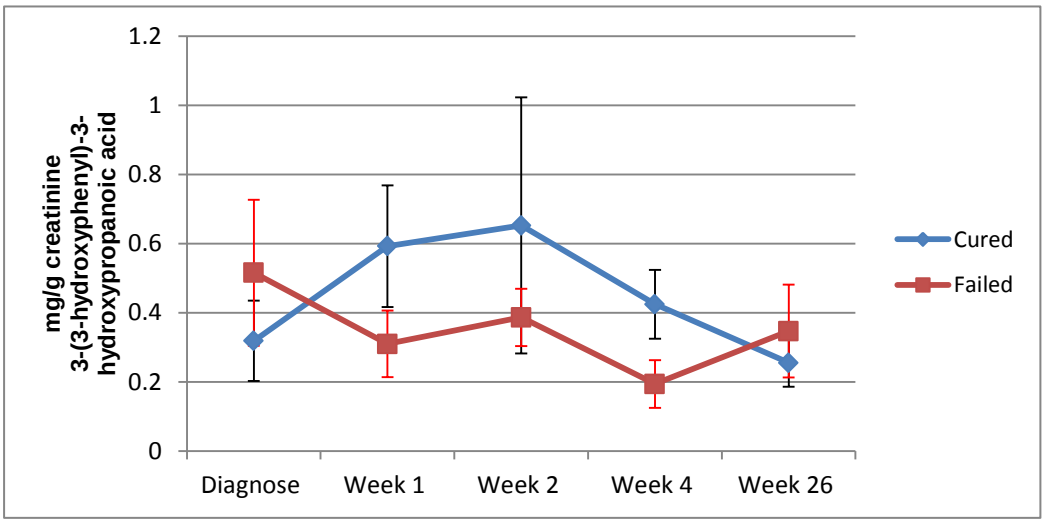
ag


mg/g creatinine Hippuric acid
80
70
60
50
40
30
20
10
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

ah


mg/g creatinine Cis-4-hydroxycyclohexanecarboxylic acid
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed

ai


mg/g creatinine 3-(3-hydroxyphenyl)-3-hydroxypropanoic acid
1.2
1
0.8
0.6
0.4
0.2
0
Diagnose
Week 1
Week 2
Week 4
Week 26
Cured
Failed