

Characterising the urinary metabolic effects of HIV/TB co-infection, using GC-MS metabolomics

C Olivier

 **orcid.org/0000-0003-4764-9054**

Dissertation accepted in partial fulfilment of the requirements
for the degree *Master of Science in Biochemistry* at
the North-West University

Supervisor: Dr L Luies

Graduation May 2024

ACKNOWLEDGEMENTS

I would like to thank the following institutions and individuals for their support and contributions to this study:

- The South African Tuberculosis Vaccine Initiative (SATVI) and Desmond Tutu HIV centre at the University of Cape Town for providing the samples and clinical information that made this study possible.
- The North-West University (NWU) for the financial support through the Master's bursary awarded to me in 2022 and 2023, allowing me to pay for my studies.
- My friends; without the encouragement, motivation and constant laughs from my friends I would not have been able to complete this year. Especially my Lab partner Charles. Thank you for being part of my journey.
- Dr Laneke Luies, my supervisor, for being the reason for the success of this study. I have learned an immense amount of knowledge and lessons from you, and I could not have asked for a better supervisor. You were willing to help at any time throughout the entire year and the support and encouragement that you have offered means the world to me. Thank you for being such a great example and inspiration.
- My family for their financial support throughout my academic pursuits.
- My mother, Pauli, I would like to thank you for being such a great example for how I should live my life and being an inspiration for the type of person that I hope to become. I appreciate your constant love, support, and belief in my dreams.
- My father, Johan, thank you for your never-ending support and love. You are such an inspiration and someone that I will continue to look up to.
- My brother, Ian, thank you for your support and encouragement. You mean so much to me and I appreciate you.

"You have brains in your head. You have feet in your shoes. You can steer yourself in any direction you choose. You're on your own. And you know what you know. And you are the guy who'll decide where to go."

- Dr. Seuss -

SUMMARY

The human immunodeficiency virus (HIV) and tuberculosis (TB) remain prominent global health threats. Individuals with HIV face increased vulnerability to various opportunistic infections due to compromised immune systems, elevating the risk of *Mycobacterium tuberculosis* (*Mtb*) infection and subsequent TB development. The co-infection of HIV and TB is particularly concerning in regions with high prevalence, such as South Africa, contributing significantly to the burden on healthcare systems. Combining resources for the treatment of co-infected patients intensifies the challenges posed by these interconnected diseases. Notably, 8% of global incident TB cases occur in people living with HIV, with southern Africa alone exceeding 50% prevalence of HIV/TB co-infection. The number of TB-diagnosed individuals aware of their HIV status has decreased to 4.2 million, underscoring gaps in testing and treatment for co-infected cases.

The choice of urine as a sample matrix for investigating HIV/TB co-infection is optimal due to its comprehensive representation of normal and pathological cellular processes, allowing for a holistic metabolic assessment. Given the goal of non-targeted metabolomics to analyse various compound classes at varying concentrations, achieving optimal extraction is crucial. Since organic acid (OA) extraction is commonly employed for urine gas chromatography mass spectrometry (GC-MS), this study firstly explored the optimised urine volume used in our current in-house OA extraction protocol (Section A, Part 1). Hereafter, we optimised a low-volume urine preparation procedure for non-targeted GC-MS. Five extraction methods (four OA extraction variations and a “direct analysis” [DA] approach) were assessed based on repeatability, metabolome coverage, and metabolite recovery (Section A, Part 2). The DA method exhibited superior repeatability, and achieved the highest metabolome coverage, detecting 91 unique metabolites from multiple compound classes comparatively. Conversely, OA methods may not be suitable for all non-targeted metabolomics applications due to their bias toward a specific compound class. In accordance, the OA methods demonstrated limitations, with lower compound recovery and a higher percentage of undetected compounds. The DA method was further improved by incorporating an additional drying step between two-step derivatization but did not benefit from urease sample pre-treatment (Section A, Part 3).

Following optimisation, the selected extraction method was used to analyse urine samples collected from healthy individuals and individuals with HIV, TB, and HIV/TB co-infection using two-dimensional gas chromatography time-of-flight mass spectrometry (GCxGC-TOFMS), aiming to elucidate the metabolic profile of co-infection (Section B). Statistical analysis identified 25 metabolites of interest that significantly differed among the four groups.

Specifically, the results revealed significant alterations across various metabolite groups, reflecting the complex metabolic impact of HIV, TB, and their co-infection. Notable shifts were observed in amino acids and related metabolites, indicating profound changes in protein biosynthesis and breakdown. These alterations are particularly pronounced in HIV-positive individuals, where increased levels of glycine and pentanedioic acid suggest adaptations in amino acid and fatty acid metabolism. In TB patients, altered phenylalanine metabolism is highlighted by elevated 3-phenyllactic acid, linked to cachexia and disrupted insulin regulation. The study also observed a unique carbohydrate metabolism in TB, marked by decreased levels of L-threonic acid and ribitol, potentially due to impaired glucose metabolism and the interplay between TB and diabetes. Co-infection with HIV and TB further amplifies these metabolic disturbances, as evidenced by the highest levels of quinolinic acid, suggesting a synergistic effect on inflammatory responses. The findings demonstrate the intricate metabolic changes occurring in HIV and TB, offering insights into their pathogenesis and potential avenues for therapeutic intervention.

Overall, this study established an improved low-volume urine preparation approach for future non-targeted urine metabolomics applications using GC-MS. Our findings contribute to advancing the field of metabolomics and enable efficient, comprehensive analysis of urinary metabolites, which could facilitate more accurate disease diagnosis or biomarker discovery. As a proof of concept, our new approach was applied to clinical samples and found that eight major metabolite categories differed significantly between the comparative groups, with notable impacts on amino acid-related metabolites and energy metabolism. Both HIV and TB affect host metabolism, altering normal metabolic pathways. Although none of the identified metabolites were disease-specific, they provided valuable insights into induced metabolic changes. The complex interplay of these disease, however, challenges the anticipated observations of HIV/TB co-infection effects. Future studies may benefit from a more targeted approach to comprehensively understand this intricate relationship.

Key words: TB; HIV; co-infection; metabolomics; urine; GCxGC-TOFMS

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
SUMMARY.....	iii
LIST OF FIGURES.....	ix
LIST OF TABLES	xiii
LIST OF ABBREVIATIONS AND SYMBOLS	xiv
CHAPTER 1: PREFACE.....	1
1.1. BACKGROUND AND MOTIVATION.....	1
1.2. AIM	2
1.3. OBJECTIVES	2
1.4. STRUCTURE OF DISSERTATION.....	3
1.5. AUTHOR CONTRIBUTIONS	4
1.6. REFERENCES	5
CHAPTER 2: EXPERIMENTAL DESIGN.....	6
2.1. ETHICAL APPROVAL	6
CHAPTER 3: SECTION A – METHOD DEVELOPMENT	8
PART 1: URINE VOLUME OPTIMISATION OF THE IN-HOUSE ORGANIC ACID EXTRACTION.....	8
3.1. LITERATURE OVERVIEW	8
3.1.1. Chemistry of Organic Acids.....	8
3.1.2. Organic Acids in Urine	8
3.1.3. GC-MS to Study Organic Acids.....	9
3.1.4. Sample Preparation	10
3.1.4.1. Organic Acid Extraction.....	10
3.1.4.2. Urease Sample Pre-Treatment.....	10
3.1.4.3. Internal Standard Selection	11
3.1.4.4. Derivatization in GC-MS: Oximation	12
3.1.4.5. Derivatization in GC-MS: Silylation.....	13
3.2. MATERIALS AND METHODS	14
3.2.1. Reagents and Chemicals.....	15
3.2.2. Spiking Solution	15
3.2.3. Internal Standard Solution.....	15
3.2.4. Sample Preparation and Ethical Approval.....	16

3.2.5.	Metabolite Extraction Methods.....	16
3.2.5.1.	Macro Volume: In-House Organic Acid Extraction Protocol	16
3.2.5.2.	Semi-Micro Volume: Assessing Small versus Large Urine Volume	16
3.2.5.3.	Full-Micro Volume: Urine and Reagent Reduction	17
3.2.6.	GC-MS Analysis	17
3.2.7.	Data Analysis.....	18
3.3.	RESULTS AND DISCUSSION.....	18
3.3.1.	Repeatability	18
3.3.2.	Metabolome Coverage.....	19
3.3.3.	Metabolite Recovery	19
3.4.	CONCLUSION.....	20
3.5.	REFERENCES	21
	PARTS 2 AND 3: METHOD DEVELOPMENT	24
3.6.	ABSTRACT.....	24
3.7.	INTRODUCTION	24
3.8.	MATERIALS AND METHODS	27
3.8.1.	Chemicals.....	27
3.8.2.	Standards	27
3.8.3.	Sample Preparation	27
3.9.	METABOLITE EXTRACTION	28
3.9.1.	Method 1 (Traditional OA Extraction)	29
3.9.2.	Method 2 (extraction solvent evaluation).....	29
3.9.3.	Method 3 (One-Step versus Two-Step Derivatization)	30
3.9.4.	Method 4 (Combination of Methods 2 and 3)	30
3.9.5.	Method 5 (DA)	31
3.9.6.	Method 6 (Additional Drying Step)	31
3.9.7.	Method 7 (Urease Pre-Treatment)	31
3.10.	GC-MS ANALYSIS	32
3.11.	DATA ANALYSIS.....	33
3.12.	RESULTS AND DISCUSSION.....	33
3.12.1.	Repeatability	33
3.12.2.	Metabolome Coverage.....	35
3.12.3.	Metabolite Recovery	36
3.12.4.	Selecting The Most Appropriate Method for Non-Targeted GC-MS Application.....	41

3.12.5.	Additional Drying Between Two-Step Derivatization.....	42
3.12.6.	Urease Sample Pre-Treatment	42
3.12.7.	Implications of The Findings	46
3.13.	CONCLUSION.....	47
3.14.	REFERENCES	47
CHAPTER 4: SECTION B – METHOD APPLICATION FOR CHARACTERISING THE URINARY METABOLIC EFFECTS OF HIV/TB CO-INFECTION.....		50
4.1.	LITERATURE REVIEW	50
4.1.1.	HIV Virology.....	50
4.1.2.	HIV Basic Pathophysiology	50
4.1.3.	TB Bacteriology	51
4.1.4.	TB Basic Pathophysiology	51
4.1.5.	The Syndemic of HIV and TB in South Africa.....	52
4.1.5.1.	WHO Goals and Beyond: Managing HIV/TB Co-infection in South Africa....	52
4.1.5.2.	The Co-Dependent Relationship Between HIV and TB.....	54
4.1.5.3.	Social and Biological Dynamics in South Africa That Contribute to HIV/TB Co-Infection.....	57
4.1.5.4.	Current Progress Supporting South Africa’s Efforts Against HIV and TB	58
4.1.5.5.	Adding Fuel to The Fire	60
4.1.5.6.	Are The “End TB” and “End HIV” Goals Out of Reach?	61
4.2.	EXPERIMENTAL DESIGN.....	63
4.3.	MATERIALS AND METHODS	64
4.3.1.	Sample Collection, Storage, Transport, and Cohort Demographics	64
4.3.2.	Reagents and Chemicals	66
4.3.3.	Urine Dipsticks.....	66
4.3.4.	Creatinine Determination	66
4.3.5.	Sample Extraction and Derivatization.....	67
4.3.6.	GCxGC-TOFMS Analysis	67
4.3.7.	Data Processing and Data Clean-Up	68
4.3.8.	Statistical Data Analyses.....	71
4.4.	RESULTS	72
4.4.1.	Machine and Analyst Repeatability	72
4.4.2.	Urine Dipsticks.....	72
4.4.3.	Univariate Testing	75
4.4.3.1.	Kruskal-Wallis Test: Statistical Comparison of All Four Groups	75

4.4.3.2. Wilcoxon Signed-Rank Test: Pairwise Statistical Comparison	76
4.4.3.3. Summary.....	78
4.4.4. Multivariate Testing.....	79
4.4.4.1. PCA and PLSDA:	79
4.4.4.2. Cross-Validation and Permutation Testing.....	81
4.5. DISCUSSION	82
4.5.1. Amino Acids and Related Metabolites.....	82
4.5.2. Carbohydrates, TCA cycle, and Energy Metabolism	84
4.5.3. Dicarboxylic Acids and Lipids.....	85
4.5.4. Gut Microbiota Imbalance	87
4.5.5. DNA Damage and Oxidative Stress	88
4.5.6. Dietary Origin.....	89
4.5.7. Exogenous Compounds.....	89
4.6. CONCLUSION.....	90
4.7. REFERENCES	91
CHAPTER 5: CONCLUSION AND FUTURE PROSPECTS	98
5.1. FINAL CONCLUSIONS	98
5.2. FUTURE PROSPECTS	99
APPENDIX A	101
APPENDIX B	102
APPENDIX C	103
APPENDIX D	104

LIST OF FIGURES

Chapter 2:

Figure 2.1: Comprehensive overview of the experimental design. This figure highlights the optimisation process in Section A, progressing from urine volume assessment to extraction method refinement. It seamlessly transitions to Section B, depicting the application of the optimized method to patient samples, offering a visual summary of the comprehensive approach in this dissertation. 7

Chapter 3:

Figure 3.1: General structure of an organic acid. In this figure, R refers to an organic group or hydrogen atom. 8

Figure 3.2: Methoximation process in GC-MS urine sample preparation. This reaction depicts the addition of methoxyamine to the carbonyl group of glucose, showcasing the protection of the carbonyl group and simplification of GC-MS analysis. 13

Figure 3.3: Reaction equation for trimethylsilylation of glucose. This reaction demonstrates the substitution of the active hydrogens in glucose with trimethylsilyl groups, a modification that reduces compound polarity and enhances volatility for GC-MS analysis. 14

Figure 3.4: Detailed experimental design for optimising urine volume in organic acid extraction. This figure outlines the experimental steps taken to determine the most effective urine volume for the extraction of organic acids using the in-house method developed. It details the variables tested, the range of urine volumes assessed, and the criteria used to evaluate the efficacy of each volume in the extraction process. Abbreviations: HCl: Hydrogen chloride; IS: internal standard; GC-MS: Gas chromatography-mass spectrometry. 14

Figure 3.5: Coefficient of Variation (CV) distribution across the in-house, semi-micro, and full-micro methods. This figure illustrates the percentage of compounds with CV values below 50% for each extraction method, providing a visual representation of method repeatability. 19

Figure 3.6: Summary of the experimental design. This study comprises a Part 2 (to evaluate five extraction methods) and Part 3 (to optimise a urine preparation procedure for non-targeted GC-MS analyses). The variations between the methods are indicated in grey sections. For Methods 1–5, six QC aliquots were used each, while Method 6 required 18 aliquots and Methods 7 needed 30 aliquots. 28

Figure 3.7: Evaluation of repeatability of the five extraction methods. (a) The percentage of metabolites below a coefficient of variation (CV) threshold < 50%, ranked as Method 5 (82.4%), Method 3 (62.1%), Method 4 (61.9%), Method 1 (54.6%), and Method 2 (36.4%). (b) Principal component analysis (PCA) scores plot which illustrates the clustering patterns among the samples based on their metabolite profiles/composition. The plot provides an overview of the overall similarity or dissimilarity among the methods' samples without directly incorporating the CV values. 35

Figure 3.8: The total ion chromatograms showing the effect of an additional drying step. Sample A (orange) was subjected to two-step derivatization while Sample B (green) was subjected to an additional drying step between two-step derivatization. Sample B showed a significant reduction in the signal of urea, which is considered advantageous. 42

Figure 3.9: The total ion chromatograms showing the effect of urease sample pre-treatment. Sample A (green) was subjected to urease pre-treatment while Sample B (orange) shows a urease-non-treated sample. A baseline offset of 2×10^7 was used for visualisation. .44

Chapter 4:

Figure 4.1: The relationship between the total deaths to deaths divided by incidence percentage, for HIV and TB in South Africa. Total TB deaths are indicated with a green line, TB deaths versus incidence percentage is indicated with a dotted blue line, total HIV deaths are indicated with a yellow line, and the HIV deaths versus incidence percentage is indicated with a dotted orange line. 55

Figure 4.2: Trends in HIV and TB mortality and incidence in South Africa (2012-2021). TB mortality as a percentage of total TB mortality for HIV-positive (orange line) and HIV-negative (blue line) subjects, with total TB mortality numbers (grey line), shown from 2012 to 2021. 57

Figure 4.3: Progress towards WHO TB reduction goals in South Africa. New TB cases (blue line) and total deaths (orange line) in South Africa, including the World Health Organization's global targets (grey line) for 2020, 2025, 2030 and 2035. The new cases targets are indicated with a yellow line. 62

Figure 4.4: Overview of the experimental design for Section B of the study. This figure summarises the workflow from sample collection to data analysis, aiming to achieve the study's objectives. 63

Figure 4.5: GC ramping temperature programme. The temperature profile of the GC run, starting at 70°C and methodically increasing to the final temperature of 300°C, outlining the precise conditions for the separation of compounds. 68

Figure 4.6: Overview of data processing and cleanup steps in metabolomics studies. This figure illustrates the essential steps of data processing and cleanup in metabolomics research. It shows the workflow from System Blank (SB) and Extraction Blank (EB) usage, through mass spectral total useful signal (MSTUS) analysis, to quality control-coefficient of variance (QC-CV) assessment. The role of LECO Corporation's ChromaTOF software in deconvolution, peak alignment, and peak identification is highlighted, showcasing the transition from raw data to analysed metabolites. 68

Figure 4.7: Principal component analysis scores plot indicating absence of batch effects. This plot visually confirms that the QC samples do not exhibit any significant between-batch or within-batch effects, demonstrating consistent instrumental response and stable column performance throughout the analysis. 70

Figure 4.8: Chromatogram overlays demonstrating machine and analyst repeatability. Here, (a) shows the consistency in chromatograms from a single QC sample injected thrice, indicating machine repeatability, while (b) displays chromatograms from QC samples across seven days, highlighting consistent analyst performance. A baseline offset of 1×10^7 and 2.5×10^7 was chosen for visualisation, respectively.....72

Figure 4.9: Urine dipsticks layout. This image (from Roche) shows the layout of the urine dipsticks used in the study, aiding in understanding the methodology and interpretation of results. 73

Figure 4.10: Kruskal-Wallis test visualisation. This figure illustrates the outcomes of the Kruskal-Wallis test, where red dots represent statistically significant metabolites, and green dots denote metabolites not found to be significant. 75

Figure 4.11: Comparative Analysis Using PCA and PLS-DA. This figure illustrates the group comparisons using PCA for inherent data structure visualisation, and PLS-DA for group classification to emphasise the discriminative features between the groups. Specifically, panel (a) shows the PCA of HIV-positive individuals compared to healthy controls, highlighting differences in their metabolic profiles. Panel (b) presents the PCA for TB-positive individuals versus healthy controls. Panel (c) illustrates the PLS-DA for HIV-positive individuals compared to healthy controls, emphasising discriminative features between these groups. Finally, panel (d) shows the PLS-DA for TB-positive individuals versus healthy controls. 80

Figure 4.12: Cross-validation test results. Here, (a) displays the cross-validation tests for the HIV-positive group versus healthy controls, and (b) shows the TB-positive group versus healthy controls. 81

Figure 4.13: Permutation test results. Permutation tests results for the (a) HIV-positive group versus healthy controls and the (b) TB-positive group versus healthy controls. 81

Figure 4.14: Metabolic alterations in TB, HIV, and HIV/TB co-infection. This highlights key metabolic pathways impacted, including amino acid, fatty acid, and carbohydrate metabolism. The interplay between TB and HIV leads to distinct changes in various metabolites, reflecting the diseases' profound effects on the body's metabolic processes. 90

LIST OF TABLES

Chapter 1:

Table 1.1 Researchers involved in the study.	5
---	---

Chapter 3:

Table 3.1: The advantages and disadvantages of urease sample pre-treatment.	11
--	----

Table 3.2: Metabolite recovery comparison for the in-house, semi-micro, and full-micro extraction methods, based on the spiking solution compounds.	20
--	----

Table 3.3: Overall performance ranking based on analytical repeatability, metabolome coverage, and recovery percentage.	20
--	----

Table 3.4: Unique compounds detected for each method, and those detected in only the organic acid extraction methods.	37
--	----

Table 3.5: Metabolite recovery of the compounds in the spiking mixture (containing pure, known standard concentrations of different class compounds) indicated for each method. The compounds are grouped per class.....	38
--	----

Table 3.6: Method rankings based on validation parameters, and an overall cumulative score of each.	41
--	----

Table 3.7: Methods rankings based on validation parameters, when comparing urease-treated, urease-nontreated, and water-treated samples.	43
---	----

Table 3.8: Method rankings based on validation parameters, when comparing heating and sonication following urease pre-treatment.	45
---	----

Chapter 4:

Table 4.1: Cohort demographics and clinical data.	65
--	----

Table 4.2: Summary of urine dipsticks results, comparing the presence and levels of various urinary markers across different sample groups.	73
--	----

Table 4.3: Significant metabolites identified by the Kruskal-Wallis test.	76
--	----

Table 4.4: Trend in the TB-positive group when compared to healthy controls.	77
---	----

Table 4.5: Significant metabolite summary.	78
---	----

LIST OF ABBREVIATIONS AND SYMBOLS

Abbreviation	Definition	Abbreviation	Definition
AIDS	Acquired Immunodeficiency Syndrome	ANOVA	one-way analysis of variance
ART	Antiretroviral therapy	BSTFA	N,O-bis(trimethylsilyl)-trifluoroacetamide
CHM	Centre for Human Metabolomics	CoA	Coenzyme A
COVID-19	Corona Virus Disease-2019	CV	Coefficient of variation
DA	Direct analysis	FADH ₂	Reduced flavin adenine dinucleotide
FDA	Federal drug administration	FDR	False discovery rate
FIND	Foundation for Innovative New Diagnostics	GABA	4-aminobutyric acid
GC	Gas Chromatograph	GC-MS	Gas Chromatograph- Mass Spectrometer
GCxGC-TOFMS	Two-dimensional gas chromatography time-of-flight mass spectrometry	HCl	Hydrogen chloride
HIV	Human immunodeficiency virus	LC	Liquid chromatography
LLE	Liquid-Liquid extraction	MSI	Metabolomics Standards Initiative
MOX-HCl	Methoxamine hydrochloride	MS	Mass Spectrometer
MSTFA	N-methyl-N-(trimethylsilyl)trifluoroacetamide	MSTUS	Mass total useful signal
<i>Mtb</i>	Mycobacterium tuberculosis	MTB/RIF	Mycobacterium tuberculosis resistance to rifampin
MTBSTFA	N-tert-butyldimethylsilyl-N-methyltrifluoroacetamide	NADH	Reduced nicotinamide adenine dinucleotide
NMR	Nuclear magnetic resonance	NSP	National Strategic Plan
NWU	North-West University	OA	Organic acids
PCA	Principle Component Analysis	PLS-DA	Partial least squares discriminant analysis
PrEP	Pre-exposure prophylaxis	QC	Quality control
QC-CV	quality control-coefficient of variance	ROS	Reactive oxygen species
SANAC	South African National AIDS Council	SATVI	South African Tuberculosis Vaccine Initiative
SCFA	short-chain fatty acid	SPE	Solid phase extraction

TB	Tuberculosis	TCA	Tricarboxylic acid cycle
TMCS	Trimethylsilyl chloride	TOF	Time Of Flight
UCT	University of Cape Town	UNAIDS	Joint United Nations Programme on HIV/AIDS
WHO	World Health Organization		

Symbol	Definition	Symbol	Definition
+	Positive	α	Alpha
-	Negative	β	Beta
$\uparrow$	Increased	μL	Microliter
$\downarrow$	Decreased	mL	Millilitre
$^{\circ}\text{C}$	Degrees Celsius	m/z	Mass to charge ratio

CHAPTER 1: PREFACE

1.1. BACKGROUND AND MOTIVATION

The human immunodeficiency virus (HIV) and tuberculosis (TB) remain significant global health threats. HIV, causing acquired immune deficiency syndrome (AIDS) (HIV/AIDS, 2021), is transmitted through sexual contact, blood contact or transfusion, and mother-to-infant transmission (World Health Organization, 2021). In 2021, there were 1.5 million new HIV infections and 680 000 AIDS-related deaths, with 27.5 million individuals globally on antiretroviral medication. Unfortunately, 10.2 million of the 37.7 million people living with HIV are not currently on treatment (Bruchfeld *et al.*, 2015).

TB, caused by *Mycobacterium tuberculosis* (*Mtb*), is transmitted through infectious aerosols usually by means of coughing and sneezing, and typically localise in the lungs (Liebenberg *et al.*, 2021; Weinrib *et al.*, 2020). Four outcomes may occur upon inhalation: immediate bacterial clearance, latent infection, active disease manifestation, or activation of latent infection (Karim & Baxter, 2010). Until 2019, TB surpassed HIV/AIDS as the leading cause of death from a single infectious agent. The global TB diagnosis rate decreased between 2020 and mid-2021, with 1.3 million TB deaths in 2020, excluding 214 000 deaths among HIV-positive individuals. Currently, 150 359 individuals receive treatment for drug-resistant *Mtb* (reduced from 2019), and 2.8 million people globally receive TB preventative treatment (Harris *et al.*, 2022).

People with HIV face increased vulnerability to opportunistic infections, including a higher risk of *Mtb* infection and TB development, making TB the most common mycobacterial disease in developing countries like South Africa (Rewari *et al.*, 2021). HIV/TB co-infection is a major health threat, especially in regions with highly prevalence (Debelu, 2022), placing a substantial burden on healthcare systems. The interconnected relationship between HIV and TB leads to severe adverse effects (Bruchfeld *et al.*, 2015; Liebenberg *et al.*, 2021), with 8% of global TB cases occurring in people living with HIV. Southern Africa alone exceeds a 50% prevalence of HIV/TB co-infection. The number of TB-diagnosed individuals aware of their HIV status has decreased to 4.2 million, revealing gaps in testing and treatment (World Health Organization, 2021).

South Africa ranks 8th globally for total TB cases, accounting for 3.3% of the total global TB cases, and 1st for HIV cases, with 7.8 million people living with HIV, constituting 13% of global cases. In 2020, Africa accounted for a major 25% of the total global TB cases, with South Africa making up a substantial 13.2% of these. Africa also reported the highest cases of HIV/TB co-infection, with South Africa contributing an astounding 50% (HIV/AIDS, 2021; World Health Organization, 2021).

While several metabolomics studies have identified distinctive metabolic associations with HIV and TB independently (Liebenberg *et al.*, 2021), comprehensive characterisation of HIV/TB co-infection is lacking. Existing studies have focused on specific aspects, for example tryptophan and kynurenine. This study aims to provide a more comprehensive understanding of HIV/TB co-infection, facilitating improved diagnostic and treatment approaches, disease monitoring, and control.

1.2. AIM

The project aims to employ a GC-MS metabolomics research approach to characterise the urinary metabolic effects of HIV/TB co-infection.

1.3. OBJECTIVES

Considering the above-mentioned aim, the objectives of this study are as follows:

- 1) Optimise a urinary metabolite extraction protocol, evaluating parameters like repeatability, metabolome coverage, and metabolite recovery for GC-MS analysis.
- 2) Extract and analyse all patient collected samples (n=115) and quality controls.
- 3) Perform data processing, clean-up, and statistical analyses to identify significant metabolite markers.
- 4) Interpret the data based on previously published literature:
 - a) Determine the impact of TB by comparing healthy controls to a TB-positive group.
 - b) Determine the impact of HIV by comparing healthy controls and HIV-positive group.
 - c) Determine the overall effect of HIV/TB co-infection by comparing an HIV/TB co-infected group to healthy controls, indicating how this corresponds with 4a and 4b — assessing whether the effect is the same or different.

1.4. STRUCTURE OF DISSERTATION

This dissertation is written in article format, to comply with the guidelines of the North-West University (NWU), Potchefstroom Campus, South Africa, for the completion of the degree Magister Scientiae (Biochemistry). Thus, the main chapters (Chapters 2 and 3) have their own literature review/introduction, materials and methods, results, discussion, conclusion, and reference list. Additionally, a summary of all findings and future prospects are provided (Chapter 5).

Chapter 1 serves as the preface and provides a succinct background and motivation for the study, outlines the aim and objectives, and delineates the basic structure of the dissertation. Additionally, it encompasses the research outputs and publications derived from the study.

Chapter 2 provides a comprehensive overview of the experimental design, offering orientation for the reader. This chapter outlines the two main sections: Section A, focused on method development and optimisation (sub-divided into Parts 1, 2, and 3), and Section B, dedicated to the application of the optimised method to patient samples.

In Chapter 3, Section A unfolds with a detailed exploration of the method development and optimisation process. Part 1 delves into the urine volume optimisation of our current in-house organic acid extraction method, whereafter Part 2 focuses on the assessment of five extraction methods in terms of repeatability, metabolome coverage, and metabolite recovery. Part 3 then investigates further optimisation of the superior method identified in Part 2. This chapter lays the foundation for the refined methodology chosen for subsequent analysis. Parts of this chapter was published (see Appendix A) and presented at an international conference (see Appendix B):

- Olivier, C., Allen, B. and Luies, L. (2023). Optimising a urinary extraction method for non-targeted GC–MS metabolomics. *Scientific Reports*, 13(1), p.17591. (<https://doi.org/10.1038/s41598-023-44690-7>)
- Olivier, C. and Luies, L. (2023). Optimising a urinary extraction method for non-targeted GC-MS metabolomics. Presented (as poster by L. Luies) at the 19th Annual Conference of the Metabolomics Society, Niagara Falls, Canada, 18–22 June 2023.

Chapter 4 (Section B) bridges the gap between method development and real-world patient samples. It delves into the application of the optimised extraction method (from Chapter 3, Part 3) to analyse urine samples collected from diverse cohorts, including healthy individuals and those with HIV, TB, and HIV/TB co-infection. This section aims to unveil the metabolic profile associated with co-infection. Parts of this chapter was published as a review (see Appendix C), while the data obtained has been drafted for journal submission:

- Olivier, C. and Luies, L. (2023). WHO Goals and Beyond: Managing HIV/TB co-infection in South Africa. *SN Comprehensive Clinical Medicine*, 5(1), p.251. (<https://doi.org/10.1007/s42399-023-01568-z>)
- Olivier, C. and Luies, L. (2023). Metabolic Insights into HIV, Tuberculosis, and Co-infection: An Untargeted Urinary Metabolomics Approach". *Drafted for intended submission*.

Chapter 5 is a comprehensive summary of the results and conclusions, in the context of the aim. In addition, future prospects are discussed.

1.5. AUTHOR CONTRIBUTIONS

The primary contributor of this dissertation is Cara Olivier, with the contributions of the co-authors to publications, collaborators, and co-workers summarised in Table 1.1. The following statement from the primary contributor and supervisor attests to their roles in the study and grants permission for the inclusion of data and conclusions in this dissertation: I affirm that my role in this study, delineated in Table 1.1, accurately reflects my actual contribution. I hereby grant my consent for the publication of this work as an integral part of the M.Sc. dissertation of Cara Olivier.



Cara Olivier



Dr. Laneke Luies

Table 1.1 Researchers involved in the study.

Co-author/Collaborator	Contribution
Cara Olivier Author/primary contributor (B.Sc. Hons. Biochemistry)	Conceptualised, planned, executed experiments, analysed data and wrote the dissertation and all associated documents and publications, together with the supervisor.
Dr Laneke Luies Co-author/supervisor (Ph.D. Biochemistry)	Supervisor: Conceptualised and co-ordinated all aspects of the study, including study design, planning, execution, writing of the dissertation and all associated documents and publications.
The South African Tuberculosis Vaccine Initiative and the Desmond Tutu HIV Centre at the Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa Collaborator	Collected and provided all serum samples and associated clinical information.
Bianca Allen Co-author (B.Sc. Hons. Biochemistry)	Assisted with some aspects of method development and data interpretation, also contributing to the related publication.

1.6. REFERENCES

Bruchfeld, J., Correia-Neves, M. & Källénus, G. 2015. Tuberculosis and HIV coinfection. *Cold Spring Harbor perspectives in medicine*, 5(7):a017871.

Debelu, B.B. 2022. Prevalence, Risk Factors and Residential Variation among HIV/TB co-infected Mortality in Amhara Region, Fenote Selam Hospital: Application of Multilevel Logistic Regression.

Harris, R.C., Quaife, M., Weerasuriya, C., Gomez, G.B., Sumner, T., Bozzani, F. & White, R.G. 2022. Cost-effectiveness of routine adolescent vaccination with an M72/AS01E-like tuberculosis vaccine in South Africa and India. *Nature Communications*, 13(1):602.

HIV/AIDS, J.U.N.P.o. 2021. *UNAIDS Data 2021*. Available from: https://www.unaids.org/en/resources/documents/2021/2021_unaids_data Date of access: 2021.

Karim, S.S.A. & Baxter, C. 2010. Microbicides & their implications in HIV prevention. *The Indian Journal of Medical Research*, 132(6):656.

Liebenberg, C., Luies, L. & Williams, A.A. 2021. Metabolomics as a Tool to Investigate HIV/TB Co-Infection. *Frontiers in Molecular Biosciences*, 8:1-19.

Rewari, B.B., Kumar, A., Mandal, P.P. & Puri, A.K. 2021. HIV TB coinfection-perspectives from India. *Expert review of respiratory medicine*, 15(7):911-930.

Weinrib, R., Browne, E., Shapley-Quinn, M., van der Straten, A., Beksinska, M., Mgodhi, N., ... Ju, S. 2020. Perspectives from young south African and Zimbabwean women on attributes of four (placebo) vaginal microbicide delivery forms. *AIDS and Behavior*, 24:637-647.

World Health Organization. 2021. *Global Tuberculosis Report*. Geneva: World Health Organization. Available from: <https://www.who.int/publications/i/item/9789240037021> Date of access: 2021.

CHAPTER 2: EXPERIMENTAL DESIGN

This dissertation comprises two main sections: Section A, divided into three parts, focuses on method development and optimisation. In Part 1, the optimal urine volume of our in-house organic acid (OA) extraction method is evaluated. Subsequently, Part 2 systematically assesses five methods based on repeatability, metabolome coverage, and metabolite recovery: An OA extraction method (Method 1) was compared with modified versions, namely the use of an additional extraction solvent (Method 2), two-step derivatization (Method 3), and a combination of Methods 2 and 3 (Method 4). A DA method (Method 5) was also tested. In Part 3, the superior method was further optimised by investigating the effects of an additional drying step between two-step derivatization (Method 6). Hereafter, urease sample pre-treatment was evaluated (Method 7) by comparing urease-treated samples to both urease-non-treated and water-treated samples. Furthermore, the effect of the heating process that urease pre-treatment requires was investigated by comparing the urease-treated to urease-non-treated samples subjected to both heating and sonication.

In Section B, the focus shifts from method development to the practical application of the optimised extraction method. This involves the analysis of urine samples collected from diverse cohorts, encompassing healthy individuals (HIV-/TB-; n=32), and those untreated for HIV (HIV+/TB-; n=9), TB (HIV-/TB+; n=41), and HIV/TB co-infection (HIV+/TB+; n=9). Quality control (QC) measures are implemented, involving the compilation of a pooled QC sample from small aliquots of each urine sample. The selected samples were extracted using the superior method identified in Section A Part 3 and analysed via GCxGC-TOFMS analysis.

The experimental design, depicted in Figure 2.1, illustrates the sequential flow of the study. It showcases the optimisation journey in Section A, from urine volume assessment to the refinement of the extraction method. Subsequently, it transitions to Section B, where the optimised method is applied to patient samples, providing a visual representation of the comprehensive approach undertaken in this dissertation.

2.1. ETHICAL APPROVAL

This study adhered to the principles outlined in the Declaration of Helsinki and International Conference on Harmonisation Guidelines. Ethical approval for the larger (broader) study was obtained from the Research Ethics Committees of NWU (reference no. NWU-00355-20-A1) and the University of Cape Town (reference no. 680/2013). Additionally, ethical approval for this specific sub-study was obtained (reference no. NWU-00355-20-A1-03). Written informed consent was obtained from all participants who willingly agreed to take part in the study.

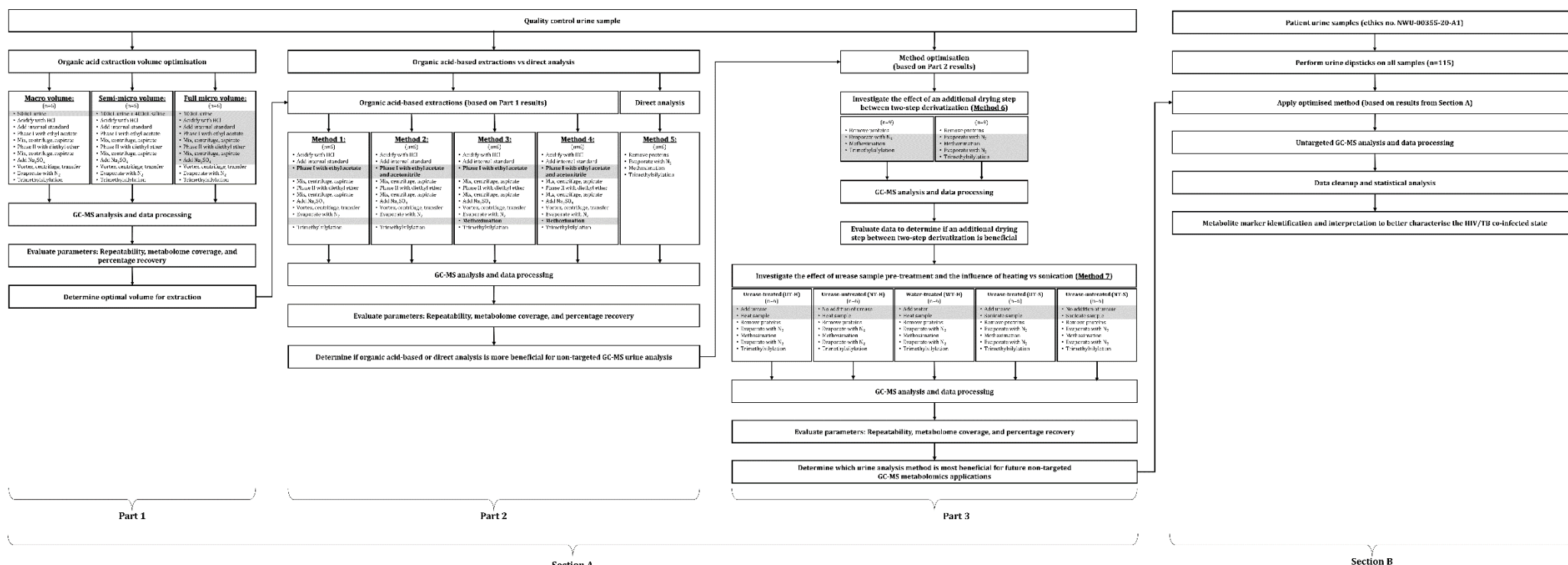


Figure 2.1: Comprehensive overview of the experimental design. This figure highlights the optimisation process in Section A, progressing from urine volume assessment to extraction method refinement. It seamlessly transitions to Section B, depicting the application of the optimized method to patient samples, offering a visual summary of the comprehensive approach in this dissertation.

CHAPTER 3: SECTION A – METHOD DEVELOPMENT

PART 1: URINE VOLUME OPTIMISATION OF THE IN-HOUSE ORGANIC ACID EXTRACTION

3.1. LITERATURE OVERVIEW

3.1.1. Chemistry of Organic Acids

Organic acids, crucial in fundamental metabolic processes, are defined by key attributes like water solubility, acidity, and the absence of primary or secondary amines (Hajati, 2018). These carboxylic acids, containing keto, hydroxyl, or other non-amino functional groups, play vital roles in metabolic pathways such as the Krebs cycle, fatty acid β -oxidation, carbohydrate metabolism, ketone body metabolism, protein metabolism and neurotransmitter turnover (Tsoukalas *et al.*, 2017). While certain nitrogen-containing compounds, such as pyroglutamate and hippurate/benzoglycine, as well as short-chain fatty acids are considered organic acids (Hajati, 2018), the general structure is defined as R-COOH, where R refers to an organic group or hydrogen atom (see Figure 3.1) (Cherrington *et al.*, 1991; Lawson *et al.*, 1976).

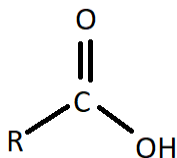


Figure 3.1: General structure of an organic acid. In this figure, R refers to an organic group or hydrogen atom.

These organic acids serve as crucial intermediate metabolites in several metabolic pathways (Tsoukalas *et al.*, 2017), reflecting the activity of major metabolic pathways in urine. Their quantification can be used to assess health, nutritional status, vitamin deficiencies, response to pharmaceutical interventions, and toxin levels in patients (Astarita & Langridge, 2013; Jones & Bennett, 2010; Shaw, 2011).

3.1.2. Organic Acids in Urine

Organic acids, abundant in nature and originating from animals, plants, and micro bacteria, are also produced and excreted in human urine (Moo-Young, 2019). These acids are found in various physiological fluids, including blood, cerebrospinal fluid, amniotic fluid, urine, and saliva, with urine being the most readily available for studying organic acids (Nordmann &

Nordmann, 1961). Human urine encompasses a multitude of organic compounds, including organic acids, making it a suitable physiological fluid for organic acid extraction procedures (Kauna-Czaplińska, 2011).

However, the use of urine samples introduces interindividual variation in excreted metabolites, influenced by external factors like diet or gender. This variation complicates the normalisation of urine metabolite concentrations and stands as the main disadvantages of this sample matrix (Du Preez & Luies, 2019). Despite this limitation, the non-invasive and easily collectable nature of urine sample (Ryan *et al.*, 2011), along with well-established protocols for collection and storage, make urine a practical choice (Du Preez & Luies, 2019). Analytically, urine presents advantages such as lower protein content, smaller size, and higher thermodynamic stability of urinary peptides and proteins, simplifying sample pre-treatment of studying organic acids (Ryan *et al.*, 2011). Although urine has its disadvantages, considering its benefits underscores its suitability as a sample medium for this study.

3.1.3. GC-MS to Study Organic Acids

Sample compositions involve mixtures of compounds and metabolites, emphasising the significance of gas chromatography (GC) in separating these components (Karasek & Clement, 2012). Mass spectrometry (MS) generates ions from compounds, and the subsequent separation based on their mass-to-charge ratio (m/z) allows for qualitative and quantitative detection of ions by their m/z and abundance (Gross, 2006). MS stands out as a method delivering meaningful data from a detector (Hübschmann, 2015). The integration of GC and MS forms a powerful technique for separating and identifying components in a complex mixture, making it ideal for biochemical analysis of samples (Hübschmann, 2015; Karasek & Clement, 2012).

GC-MS is recognised as the “gold standard” method in metabolomics studies, supported by extensive research and established methods (Fiehn, 2016). It is particularly preferred for studying small molecule metabolites and primary metabolites such as amino acids and organic acids (Wajner *et al.*, 2009). GC-MS boasts several advantages, including the availability of commercial spectral libraries available, consistent results across different laboratories, excellent chromatographic resolution, and a wealth of literature and experience (Christou *et al.*, 2014).

A notable improvement for GC-MS involves coupling the system to a “time-of-flight” (TOF) mass spectrometer. In this setup, ion pulses are produced, and the separation of ions with different m/z ratios occur based on their flight times to the detector. Introducing a TOFMS setup significantly broadens its applications (Mamyrin, 2001).

3.1.4. Sample Preparation

3.1.4.1. Organic Acid Extraction

Sample preparation is a critical step in any analytical procedure, involving the isolation and pre-concentration of the analyte of interest (Maya *et al.*, 2018). Successful sample preparation is essential for developing reliable methods that achieve high sensitivity, accuracy, precision, and recovery in results (Płotka-Wasyłka *et al.*, 2015).

In this experimental procedure, organic acids are the analytes of interest. Before analyses, organic acids must be extracted to render the sample matrix suitable for GC-MS analysis. Two main methods for extracting organic acids are liquid-liquid extraction (LLE) and solid-phase extraction (SPE), with LLE being the more common technique (Kumari *et al.*, 2016). Each method has its advantages and disadvantages, and the choice depends on specific research needs and available equipment.

SPE involves an aqueous sample passing through an immobilised phase, subsequently extracted by organic solvents based on the selective distribution of the analyte between the solid packing material and the liquid mobile phase. The workflow consists of column preparation, sample loading, column post-wash, and sample desorption (Buszewski & Szultka, 2012). Despite the high costs and complexity associated with SPE, its benefits include analyte concentration, matrix removal, and automation potential (Stone, 2017). Compared to LLE, SPE isolates more metabolites, particularly organic acids analysed by GC-MS, uses less organic solvents, offers clean extracts with superior recoveries compared to other extraction methods, and is faster, making it suitable for large sample sets (Kumari *et al.*, 2016; Verhaeghe *et al.*, 1988). However, caution is advised when working with small sample amounts due to increased time requirements.

LLE employs water-immiscible organic solvents to induce analyte partitioning between aqueous samples and the added organic solvent (Soares da Silva Burato *et al.*, 2020). While LLE is cost-effective and straightforward (Kumari *et al.*, 2016), it requires larger volumes of organic solvents and takes more time than SPE. The method is also criticised for the toxicity of commonly used extractants, posing environmental and health hazards (Chen *et al.*, 2008). Despite these disadvantages, LLE's simplicity, cost-effectiveness, and suitability for smaller sample amounts (Kumari *et al.*, 2016) make it a viable choice for this experiment.

3.1.4.2. Urease Sample Pre-Treatment

The enzyme urease has been employed since 1964 to remove excess urea from urine samples prior to GC-MS analysis, as first reported by Wells *et al.* (1964) (Webb-Robertson *et al.*, 2014).

This step is crucial in metabolomics studies to address chromatographic interference caused by high urea levels, which can overwhelm the chromatogram and hinder identification of metabolites eluting at similar times as urea (Chan *et al.*, 2011). Additionally, urea has been known to interfere with derivatization steps, leading to incomplete chemical transformation (Webb-Robertson *et al.*, 2014). While urease offers advantages, it is essential to consider potential disadvantages, such as unwanted chemical transformations and secondary enzymatic reactions (Kind *et al.*, 2007). Table 3.1 summarises the advantages and disadvantages of using urease sample pre-treatment in metabolomics studies. These insights are essential for understanding the potential impact of urease on the analysis of urine samples for GC-MS. The references provided offer a foundation for further exploration and validation of these findings.

Table 3.1: The advantages and disadvantages of urease sample pre-treatment.

Advantages of using urease	References
Prevents interference of urea with chemical derivatization	Webb-Robertson <i>et al.</i> (2014)
Prevents chromatographic column overloading	Palmas <i>et al.</i> (2018)
Prevents peak distortions	Palmas <i>et al.</i> (2018)
Prevents the chances of co-eluting peaks	Palmas <i>et al.</i> (2018)
Lowers coefficient of variation (CV) values compared to untreated samples	Webb-Robertson <i>et al.</i> (2014)
Minimises severe disturbances of chromatographic analysis	Kim <i>et al.</i> (2020)
Disadvantages of using urease	References
Possible alterations of metabolic profiles	Kim <i>et al.</i> (2020)
Conversion of creatine to creatinine causing unwanted peak formation	Matsumoto and Kuhara (1996) Shoemaker and Elliott (1991)
Contaminants such as myoinositol and citrate are introduced	Shoemaker and Elliott (1991) Clements Jr and Starnes (1975)
Decreased abundances for aconitate, acotinate, ascorbate, citrate, glycerol, hypoxanthine, succinate, and tyrosine reported	Kind <i>et al.</i> (2007)
Certain metabolites are not detected in urease-treated samples due to possible repression, such as glycerate, hippurate, pyroglutamate, and phosphate	Webb-Robertson <i>et al.</i> (2014)

3.1.4.3. Internal Standard Selection

An internal standard is a compound added to every sample before analysis, absent in the samples naturally, with chemical properties similar to those being analysed (Khodadadi & Pourfarzam, 2020). Internal standards ensure control over sample-based variations, including collection, storage, handling, extraction, and analytical programme-generated variation (Misra, 2020). Various types of internal standards exist, with this project employing a spike-in standard

(namely 3-phenylbutyric acid), added from an unrelated source and differing from the analyte class of interest. Other types of internal standards include isotopologue standards (deuterated or ^{13}C labelled analytes of interest), stable isotope-labelled cells or extracts, and retention time markers — all used for normalisation purposes (Misra, 2020). Incorporating internal standards in GC-MS metabolomics is common, employing either structure-related or stable isotopically labelled internal standard compounds (Jiang *et al.*, 2017).

Internal standards are essential for quantifying analyte concentrations based on the retention time (McDonald *et al.*, 2006). To determine the concentration of any metabolite in the sample, it is compared to the internal standard (see equation 3.1). This ratio, known as the response factor, is proportional to the compound's concentration. Ensuring consistency, the method must distinguish between the analyte of interest and the internal standard, accounting for similar sample losses after their addition (Rozio *et al.*, 2019).

$$\text{Concentration} = \frac{\text{Area of compound}}{\text{Area of internal standard}} \times \text{concentration of internal standard} \quad (3.1)$$

3.1.4.4. Derivatization in GC-MS: Oximation

Biological urine samples often contain metabolites that are non-volatile and thermally unstable, which are unsuitable for GC-MS analysis that requires volatilised compounds. To align with this requirement, derivatization of urine samples is imperative (Khodadadi & Pourfarzam, 2020). A two-step derivatization, including oximation followed by silylation, is commonly employed (Chan *et al.*, 2011), with oximation being a pivotal preparatory step (Christou *et al.*, 2014). Oximation primarily aims to simplify the chromatogram by minimising carbohydrate peaks and preserving structural integrity of α -keto acids and aldehydes against degradation (Khodadadi & Pourfarzam, 2020). Carbohydrates present in urine can form multiple peaks due to their cyclic and acyclic forms during silylation, thereby complicating the chromatogram and hindering metabolite identification (Megías-Pérez *et al.*, 2019). Oximation, therefore, is a protective step before silylation; it protects α -keto acids and aldehydes from decarboxylation and mitigates keto-enol tautomerism, thus streamlining the number of peaks and enhancing metabolite stability (Moros *et al.*, 2017).

For oximation, the choice of reagents includes hydroxylamines and alkoxyamines (Fiehn, 2016), with this project opting for methoxyamine hydrochloride ($\text{CH}_3\text{ONH}_2 \cdot \text{HCl}$) due to its straightforward chemistry and identifiable features. Methoxyamine, being an alkoxyamine (Khodadadi & Pourfarzam, 2020), serves to block the carbonyl group in sugars, inhibiting cycling hemiacetal formation and consequently multiple chromatographic peaks. It also preserves α -keto acids from decarboxylation processes (Moros *et al.*, 2017). Figure 3.2

illustrates the methoximation of glucose, wherein methoxyamine is added to the carbonyl group of glucose. This reaction shields the carbonyl group from undesirable side reactions and simplifies future analyses.

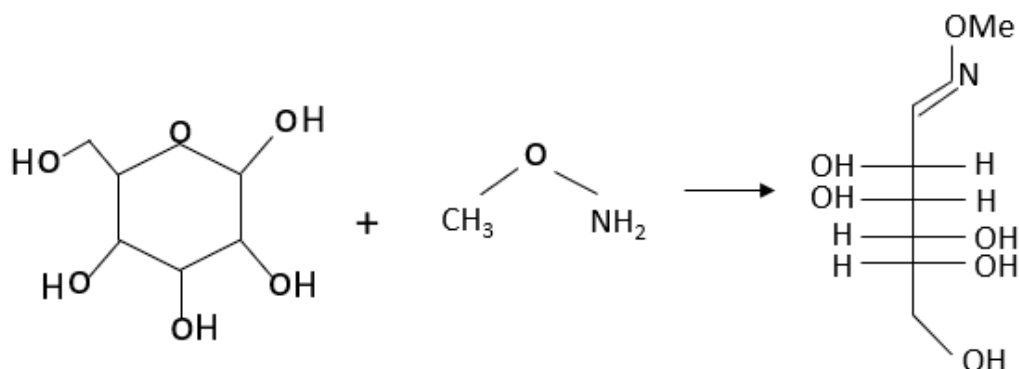


Figure 3.2: Methoximation process in GC-MS urine sample preparation. This reaction depicts the addition of methoxyamine to the carbonyl group of glucose, showcasing the protection of the carbonyl group and simplification of GC-MS analysis. Adapted from Yi et al. (2014).

3.1.4.5. Derivatization in GC-MS: Silylation

Following oximation, the second critical step in the derivatization method for GC-MS analysis is silylation. This step involves the substitution of active hydrogens with a silyl group in molecules containing such hydrogens. Active hydrogens can be found in a variety of functional groups, including acids, alcohols, enolizable ketones, amides, thiols, amines, and aldehydes. The introduction of a silyl group diminishes the polarity and hydrogen bonding potential of these compounds, thereby enhancing their volatility and making them amenable for GC-MS analysis (Khodadadi & Pourfarzam, 2020). Common silylation reagents cited in the literature include MSTFA (N-methyl-N-(trimethylsilyl)trifluoroacetamide), MTBSTFA (N-tert-butyl-dimethylsilyl-N-methyltrifluoroacetamide) and BSTFA (N,O-bis(trimethylsilyl)trifluoroacetamide), each with distinct advantages and disadvantages. A catalyst, typically 1% trimethylchlorosilane (TMCS), is often added to enhance the efficacy of these silylation agents (Jiye et al., 2005; Kanani et al., 2008; Pasikanti et al., 2008). For this project, BSTFA was selected due to its superior derivatization efficiency for carboxylic acids, despite being less volatile compared to MSTFA (Zhu et al., 2017).

Figure 3.3 illustrates the trimethylsilylation of glucose, a representative reaction wherein the active hydrogens of the glucose molecule are substituted with trimethylsilyl groups, facilitating the analysis of such derivatives in GC-MS.

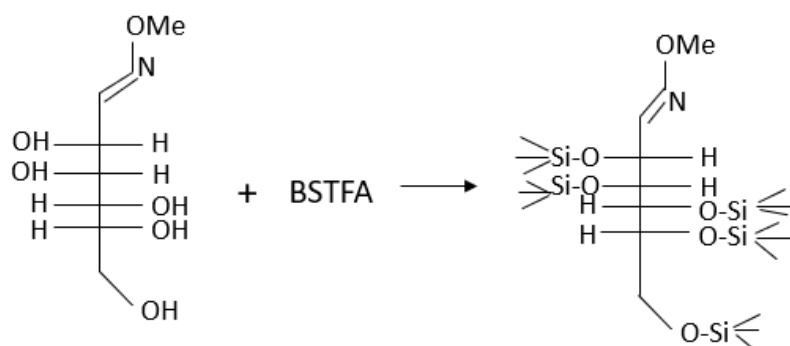


Figure 3.3: Reaction equation for trimethylsilylation of glucose. This reaction demonstrates the substitution of the active hydrogens in glucose with trimethylsilyl groups, a modification that reduces compound polarity and enhances volatility for GC-MS analysis. Adapted from Yi et al. (2014).

3.2. MATERIALS AND METHODS

While a thorough overview of this study's experimental approach is delineated in Chapter 2, a focused depiction of the experimental design specific to Section A's Part 1 is illustrated in Figure 3.4. This segment of the investigation assesses the optimal urine volume necessary for our established in-house organic acid (OA) extraction method.

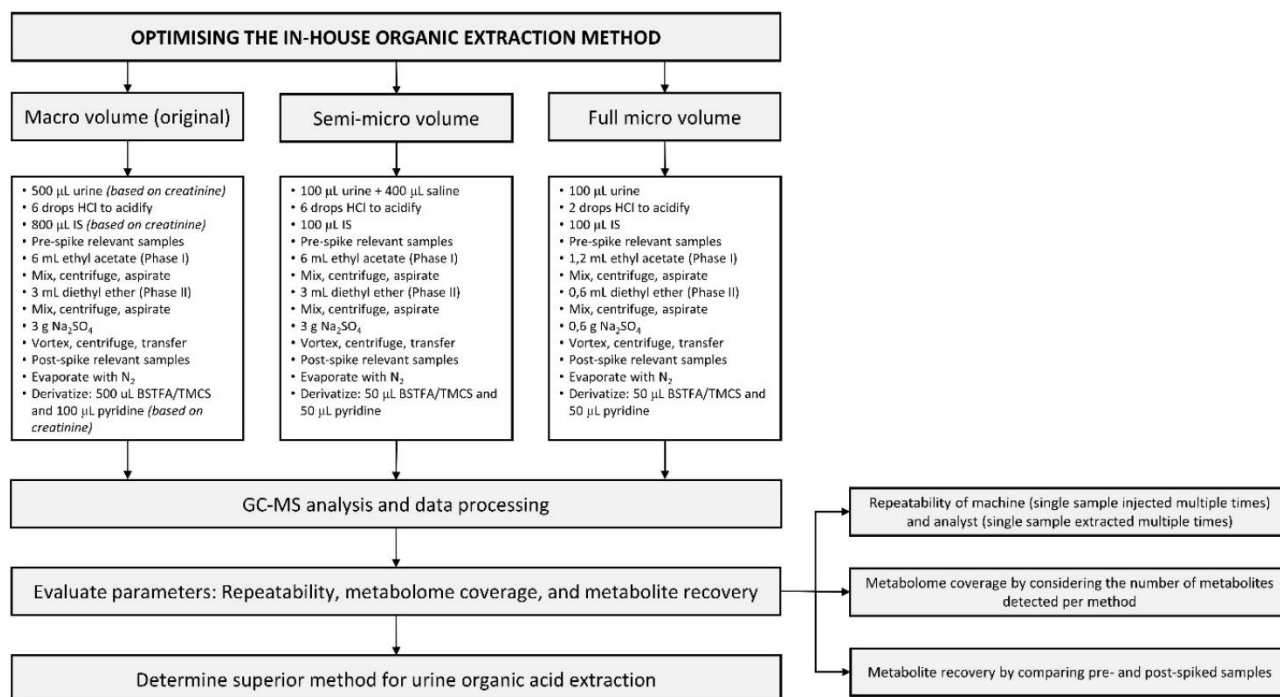


Figure 3.4: Detailed experimental design for optimising urine volume in organic acid extraction. This figure outlines the experimental steps taken to determine the most effective urine volume for the extraction of organic acids using the in-house method developed. It details the variables tested, the range of urine volumes assessed, and the criteria used to evaluate the efficacy of each volume in the extraction process. Abbreviations: HCl: Hydrogen chloride; IS: internal standard; GC-MS: Gas chromatography-mass spectrometry.

3.2.1. Reagents and Chemicals

Organic acid extraction and derivatization methods were performed on the samples using the following reagents: 3-phenylbutyric acid as well as methoxyamine hydrochloride (MOX-HCl), sodium sulphate (Na_2SO_4), urease and BSTFA with 1% TMCS. These reagents were acquired from Sigma Aldrich (St. Louis, Missouri, USA). Other chemicals used were hydrochloric acid (HCl), ethyl acetate, and diethyl ether from Burdick and Jackson brand (Honeywell International Inc., Muskegon, USA). Pyridine served as a solvent for MOX-HCl and was purchased from Merck (Darmstadt, Germany).

3.2.2. Spiking Solution

Prior to the execution of experimental assays, a spiking solution was meticulously prepared with exact quantities of analytes from diverse compound classes. The purpose of this solution is to evaluate the percentage recovery of known analytes post-sample processing. To ensure a comprehensive representation of different metabolite types, two distinct compound classes were included in the spiking mixture: amino acids (4-aminobutyric acid [GABA], L-phenylalanine, L-tryptophan), and organic acids (ascorbic acid, citric acid, succinic acid). Each compound was precisely weighed to 25 mg and transferred into separate 50 mL volumetric flasks, which was then filled to volume with milliQ water. The amino acids, being in liquid form, were measured and added distinctly. Hereafter, 5 mL from each prepared compound solution and the liquid amino acids were combined in a 50 mL beaker and diluted to the mark with milliQ water to achieve a uniform spiking solution concentration of 50 ppm.

3.2.3. Internal Standard Solution

3-Phenylbutyric acid was used as internal standard due to its non-occurrence in normal urine samples and absence in known pathological conditions. Additionally, it has the advantage of eluting near the middle of the organic acid chromatographic profile, and has a reduced likelihood of co-eluting with other organic acids, thereby simplifying the analysis (Reinecke *et al.*, 2012). To prepare the internal standard solution, 10 mg of 3-phenylbutyric acid was accurately weighed and dissolved in a volumetric flask, which was filled to volume with milliQ water. For this, 12.5 mL was transferred to a 50 mL beaker and filled to volume with milliQ water to achieve a final concentration of 50 ppm. This standardisation ensures that any variations in analyte recovery during sample processing can be monitored and corrected.

3.2.4. Sample Preparation and Ethical Approval

This study utilised a single quality control (QC) sample, which has been ethically approved, as outlined in Chapter 2. This QC sample was a composite of healthy patient samples, amalgamated to serve as a benchmark for the analytical procedures.

3.2.5. Metabolite Extraction Methods

3.2.5.1. Macro Volume: In-House Organic Acid Extraction Protocol

To each urine sample ($n=6$; 500 μL), six drops of 5 N HCl were added. This was followed by the introduction of 100 μL of internal standard solution (50 ppm) using a Hamilton syringe for precise delivery. Hereafter, half of the samples ($n=3$) were fortified with a 50 μL known spiking solution (referred to as the “pre-spiked samples”). Thereafter, 6 mL of ethyl acetate was added to each sample to facilitate phase separation. Then samples were uniformly mixed for 30 minutes on a rotary mixer to ensure thorough interaction. After mixing, the samples were centrifuged at 845 $\times g$ for 5 minutes to separate the phases, after which the top organic layer was carefully extracted using a glass Pasteur pipette. To the remaining aqueous layer, 3 mL of diethyl ether was added, mixed, and again centrifuged under the same conditions to collect the resultant organic phase, which was then combined with the first extract. Anhydrous Na_2SO_4 , approximately 3 g, was added into each sample to remove residual moisture, followed by vortex mixing and a final centrifugation. The clarified organic phase was transferred to a new tube. At this juncture, the remaining samples ($n=3$) were spiked (referred to as the “post-spike samples”). Each sample was then dried under a gentle stream of nitrogen gas at 37°C for about one hour, to ensure complete evaporation of the solvent (Luies & Loots, 2016).

For derivatization, 50 μL of BSTFA with 1% TMCS and 50 μL of pyridine were added to the dried samples using a Hamilton syringe for precision. The samples were sealed, incubated at 60°C for 60 minutes, and then allowed to cool to room temperature. Finally, the derivatized samples were transferred to GC sample vials equipped with inserts and securely capped for analysis.

3.2.5.2. Semi-Micro Volume: Assessing Small versus Large Urine Volume

In this part of the investigation, the protocol was adjusted to use a “semi-micro volume” approach. The urine volume was set to 100 μL and made up to 500 μL with 9% saline solution to mimic the larger volume typically used in the standard in-house protocol. The rest of the extraction procedure was carried out as delineated in Section 3.2.5.1. The objective was to evaluate whether the smaller urine volume, once adjusted to the larger volume with saline, would affect the extraction and analysis outcomes compared to the “large volume” method.

3.2.5.3. Full-Micro Volume: Urine and Reagent Reduction

In the “full-micro volume” method, the experiment was designed to function with a minimal volume of urine, set at 100 μL . Accordingly, the volumes of all reagents were reduced to maintain the same ratio as the standard protocol. Specifically, two drops of 5 N HCl were added to the 100 μL of urine, and this was followed by the addition of 20 μL internal standard to each sample and 50 μL of the spiking solution to a subset of three samples. For phase separation, 1.2 mL of ethyl acetate and subsequently 0.6 mL of diethyl ether were used. After phase separation, 0.6 g of anhydrous Na_2SO_4 was added, whereafter the remaining samples were spiked accordingly. The volumes of the BSTFA and pyridine for derivatization were kept the same as in the larger volume protocols to ensure consistency in the final analysis step.

The practice of reducing the volume of urine in extraction processes is referred to as “miniaturisation”. This technique involves diminishing the volume of the extraction phase to streamline experimental procedures, enhances selectivity, and improves the uptake of analytes, thus facilitating the extraction of compounds with a wide range of polarities (Soares da Silva Burato *et al.*, 2020). Several studies have corroborated the efficacy of miniaturisation, particularly in the context of organic acid LLE, demonstrating that it can lead to improved analytical results (Nakagawa *et al.*, 2010).

3.2.6. GC-MS Analysis

The analysis of samples was conducted in a randomised sequence within a single analytical batch. They were organised into the autosampler of a coupled GC-TOFMS system, utilising an Agilent 7890 Gas Chromatograph and a Time-of-Flight Mass Spectrometer provided by LECO Africa. Each sample (1 μL) was introduced into the system with a split ratio of 1:10, employing ultra-pure helium as the carrier gas with a flow rate of 1.4 mL/min. A Restek Rxi-5MS capillary column was employed for the chromatographic resolution of the analytes (with dimensions of 28.2 m length, 250 μm internal diameter, and 0.25 μm film coating). The temperature parameters were rigorously controlled: 250°C for the inlet, 225°C for the transfer line, and 200°C for the ionisation source. The thermal gradient of the oven initiated at a baseline of 50°C, maintained for 1 minute, followed by a series of programmed increases: first to 100°C at a rate of 5°C/min, then to 160°C at 10°C/min, escalating to 230°C at 13°C/min, and culminating at 300°C at 20°C/min, where it remained for an additional 2 minutes. The entire duration of each analytical run was approximately 24 minutes. A solvent delay period of 480 s preceded the recording of mass spectra to prevent solvent peaks from saturating the detector, though the corresponding time was accounted for in the chromatogram's temporal axis. Mass spectral data was acquired within the m/z range of 50–800 at a frequency of 20

spectra per second, with a filament bias set at 70 eV and detector voltage fine-tuned with a 50 V offset.

3.2.7. Data Analysis

Data processing was performed using the Leco® ChromaTOF software (version 4.72.0.0), which facilitated MS deconvolution, with a signal-to-noise threshold of 100 and requiring at least three consecutive peaks to qualify for detection. Compound identification was conducted with a minimum similarity criterion of 60% for database matches. Following this, peak alignment was conducted to ensure consistency across samples. The processed data was initially output in a .csv file format and then converted into an Excel Workbook (.xlsx) for further manipulation. Normalisation procedures were applied first against the internal standard and thereafter adjusted based on creatinine levels to account for sample dilution variability. Using Excel, parameters such as analytical repeatability, metabolome coverage, and metabolite recovery were evaluated and quantified.

3.3. RESULTS AND DISCUSSION

The efficiency and effectiveness of using different urine volumes for the extraction of organic acids were evaluated, with specific attention to analytical repeatability, metabolome coverage, and metabolite recovery.

3.3.1. Repeatability

Analytical repeatability was assessed by repeatedly extracting a single sample and injecting each extract onto the GC-MS once. The coefficient of variation (CV) was determined for all compounds detected in each method. The CV, as a percentage, delineates the ratio of the standard deviation to the mean. A lower CV indicates higher method precision. Although a CV of less than 20% is the benchmark for targeted GC-MS analysis as per Federal Drug Administration (FDA) guidelines, for non-targeted metabolomics, a CV of less than 50% is deemed acceptable due to the broader range of compounds analysed. Figure 3.5 highlights the CV distribution across the three methods, where the full-micro method exhibited a superior repeatability with 85.14% of compounds falling below the 50% CV threshold, outperforming the semi-micro method (27.16%) and the in-house method (20%).

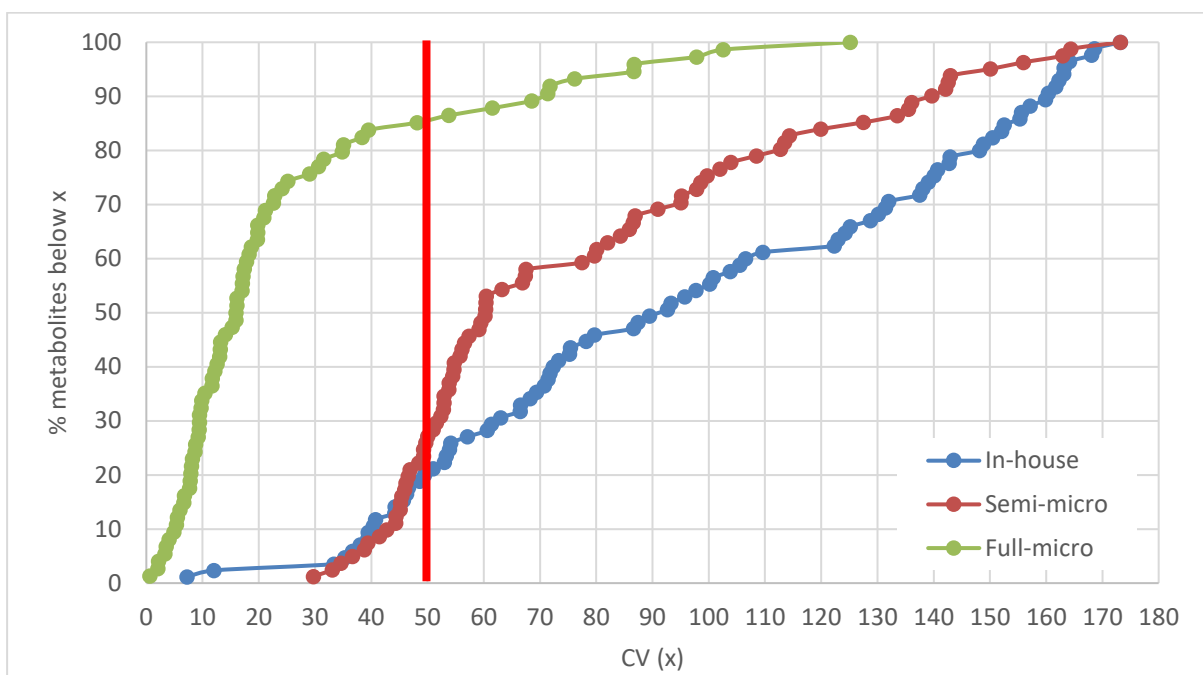


Figure 3.5: Coefficient of Variation (CV) distribution across the in-house, semi-micro, and full-micro methods. This figure illustrates the percentage of compounds with CV values below 50% for each extraction method, providing a visual representation of method repeatability.

3.3.2. Metabolome Coverage

Metabolome coverage was evaluated by considering the total number of compounds identified in each method. Out of 110 compounds detected across all methods, the full-micro method missed only 13.63%, indicating its heightened sensitivity. The in-house method failed to detect 22.73% of the compounds, while the semi-micro method missed 26.36%. Furthermore, the in-house method uniquely identified two compounds, an unknown and 2,5-di-tert-butylbenzene-1,4-diol. The semi-micro method discovered four unique compounds, including two unknowns, tridecane, and 1-hexadecanol. Notably, the full-micro method detected nine unique compounds, encompassing one unknown, two sugars (the annotation of which cannot be certain with a GC system), two compounds from the spiking solution, oxalic acid, 3-hexenedioic acid, 5-oxoproline, and 2-hydroxy-3-methylvaleric acid, demonstrating its broad-spectrum detection capability.

3.3.3. Metabolite Recovery

The metabolite recovery was measured by adding a spiking solution to the samples at the start and end of each method. As shown in Table 3.2, the semi-micro method achieved an average recovery of 43.8%, predominantly recovering succinic acid and citric acid. The in-house method recovered these acids at much lower rates, with an overall recovery of 5.2%. The full-micro method, despite detecting a broader range of spiking compounds, had the lowest

average recovery at 4.6%. The lower recovery of amino acids can be attributed to the inherent inefficiency of organic acid extraction methods in recovering such compounds.

Table 3.2: Metabolite recovery comparison for the in-house, semi-micro, and full-micro extraction methods, based on the spiking solution compounds.

Compounds	In-house method			Semi-micro method			Full-micro method		
	Pre-spike average	Post-spike average	%	Pre-spike average	Post-spike average	%	Pre-spike average	Post-spike average	%
GABA	0.0	3.2	0.0	0.0	0.7	0.00	0.0	1.6	1.4
Ascorbic acid	0.0	0.1	0.0	0.0	0.0	0.00	0.0	0.0	0.0
Succinic acid	3.2	15.4	20.8	8.8	4.2	210.9	0.7	8.8	7.4
Citric acid	1.4	12.8	10.7	1.8	3.5	52.1	0.8	6.1	12.8
Tryptophan	0.0	3.7	0.0	0.0	1.5	0.0	0.0	0.1	0.0
Phenylalanine	0.0	26.3	0.0	0.0	6.4	0.0	0.7	11.5	6.2
Average recovery %			5.2			43.8			4.6

3.4. CONCLUSION

Upon reviewing the validation parameters, each method was ranked based on its overall performance, and a cumulative score was calculated (see Table 3.3). Since a top ranking would result in a lower cumulative score, the method with the lowest cumulative score should theoretically be the superior method.

Table 3.3: Overall performance ranking based on analytical repeatability, metabolome coverage, and recovery percentage.

Method \ Parameter	Repeatability	Metabolome coverage	Percentage recovery	Cumulative score
In-house	3	2	2	7
Semi-micro	2	3	1	6
Full-micro	1	1	3	5

The full-micro method ranked highest in repeatability and metabolome coverage, while the semi-micro method led in metabolite recovery. Despite the apparent cumulative score advantage of the full-micro method, practical difficulties during its implementation warrant a more nuanced approach. Specifically, the manipulation of small volumes required in the full-micro method introduced significant challenges, such as the precise transfer of organic phases, which was problematic. The small volume of Na₂SO₄ added during the full-micro method presented an issue during decanting — a glass Pasteur pipette was needed for

accurate transfer of the organic phases rather than the free-hand decanting at this stage for the other two methods. Without the aid of automation, working with such small volumes could lead to errors affecting recovery rates.

Consequently, the recommendation for future GC-MS applications leans toward the semi-micro method. It offers a balance between operational practicality and analytical performance, mitigating the risks associated with manual handling of micro volumes. With future advancements in automation and method validation, the full-micro method's potential could be fully realised.

3.5. REFERENCES

- Astarita, G. & Langridge, J. 2013. An emerging role for metabolomics in nutrition science. *Lifestyle Genomics*, 6(4-5):181-200.
- Buszewski, B. & Szultka, M. 2012. Past, present, and future of solid phase extraction: a review. *Critical Reviews in Analytical Chemistry*, 42(3):198-213.
- Chan, E.C.Y., Pasikanti, K.K. & Nicholson, J.K. 2011. Global urinary metabolic profiling procedures using gas chromatography–mass spectrometry. *Nature protocols*, 6(10):1483-1499.
- Chen, Y., Guo, Z., Wang, X. & Qiu, C. 2008. Sample preparation. *Journal of Chromatography A*, 1184(1-2):191-219.
- Cherrington, C., Hinton, M., Mead, G. & Chopra, I. 1991. Organic acids: chemistry, antibacterial activity and practical applications. *Advances in microbial physiology*, 32:87-108.
- Christou, C., Gika, H.G., Raikos, N. & Theodoridis, G. 2014. GC-MS analysis of organic acids in human urine in clinical settings: a study of derivatization and other analytical parameters. *Journal of Chromatography B*, 964:195-201.
- Clements Jr, R.S. & Starnes, W.R. 1975. An improved method for the determination of urinary myoinositol by gas-liquid chromatography. *Biochemical medicine*, 12(2):200-204.
- Du Preez, I. & Luies, L. 2019. The application of metabolomics toward pulmonary tuberculosis research. *Tuberculosis*, 115:126-139.
- Fiehn, O. 2016. Metabolomics by gas chromatography–mass spectrometry: Combined targeted and untargeted profiling. *Current protocols in molecular biology*, 114(1):30.34. 31-30.34. 32.
- Gross, J.H. 2006. *Mass spectrometry: a textbook*. Springer Science & Business Media.
- Hajati, H. 2018. Application of organic acids in poultry nutrition. *Int. J. Avian Wildl. Biol*, 3:324-329.
- Hübschmann, H.-J. 2015. *Handbook of GC-MS: fundamentals and applications*. John Wiley & Sons.
- Jiang, R., Jiao, Y., Zhang, P., Liu, Y., Wang, X., Huang, Y., ... Xu, F. 2017. Twin derivatization strategy for high-coverage quantification of free fatty acids by liquid chromatography–tandem mass spectrometry. *Analytical chemistry*, 89(22):12223-12230.
- Jiye, A., Trygg, J., Gullberg, J., Johansson, A.I., Jonsson, P., Antti, H., ... Moritz, T. 2005. Extraction and GC/MS analysis of the human blood plasma metabolome. *Anal Chem*, 77(24):8086-8094.
- Jones, P.M. & Bennett, M.J. 2010. Urine organic acid analysis for inherited metabolic disease using gas chromatography-mass spectrometry. In. *Clinical applications of mass spectrometry*: Springer. pp. 423-431.
- Kanani, H., Chrysanthopoulos, P.K. & Klapa, M.I. 2008. Standardizing gc–ms metabolomics. *Journal of Chromatography B*, 871(2):191-201.

- Karasek, F.W. & Clement, R.E. 2012. *Basic gas chromatography-mass spectrometry: principles and techniques*. Elsevier.
- Kauna-Czaplińska, J. 2011. Current applications of gas chromatography/mass spectrometry in the study of organic acids in urine. *Critical Reviews in Analytical Chemistry*, 41(2):114-123.
- Khodadadi, M. & Pourfarzam, M. 2020. A review of strategies for untargeted urinary metabolomic analysis using gas chromatography–mass spectrometry. *Metabolomics*, 16:1-14.
- Kim, J., Ahn, J.K., Cheong, Y.E., Lee, S.-J., Cha, H.-S. & Kim, K.H. 2020. Systematic re-evaluation of the long-used standard protocol of urease-dependent metabolome sample preparation. *Plos one*, 15(3):e0230072.
- Kind, T., Tolstikov, V., Fiehn, O. & Weiss, R.H. 2007. A comprehensive urinary metabolomic approach for identifying kidney cancer. *Analytical biochemistry*, 363(2):185-195.
- Kumari, C., Varughese, B., Ramji, S. & Kapoor, S. 2016. Liquid–liquid extraction and solid phase extraction for urinary organic acids: a comparative study from a resource constraint setting. *Indian Journal of Clinical Biochemistry*, 31(4):414-422.
- Lawson, A.M., Chalmers, R.A. & Watts, R. 1976. Urinary organic acids in man. I. Normal patterns. *Clinical chemistry*, 22(8):1283-1287.
- Luies, L. & Loots, D.T. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.
- Mamyrin, B. 2001. Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry*, 206(3):251-266.
- Matsumoto, I. & Kuhara, T. 1996. A new chemical diagnostic method for inborn errors of metabolism by mass spectrometry—rapid, practical, and simultaneous urinary metabolites analysis. *Mass spectrometry reviews*, 15(1):43-57.
- Maya, F., Palomino Cabello, C., Ghani, M., Turnes Palomino, G. & Cerdà, V. 2018. Emerging materials for sample preparation. *Journal of separation science*, 41(1):262-287.
- McDonald, T.J., Wang, B., McDonald, S.J. & Brooks, J.M. 2006. Quantitative determination of aromatic hydrocarbons using selected ion monitoring gas chromatography/mass spectrometry. *TDI-Brooks International, College Station, TX*,
- Megías-Pérez, R., Shevchuk, A., Zemedie, Y. & Kuhnert, N. 2019. Characterization of commercial green tea leaves by the analysis of low molecular weight carbohydrates and other quality indicators. *Food chemistry*, 290:159-167.
- Misra, B.B. 2020. Data normalization strategies in metabolomics: Current challenges, approaches, and tools. *European Journal of Mass Spectrometry*, 26(3):165-174.
- Moo-Young, M. 2019. *Comprehensive biotechnology*. Elsevier.
- Moros, G., Chatziioannou, A.C., Gika, H.G., Raikos, N. & Theodoridis, G. 2017. Investigation of the derivatization conditions for GC–MS metabolomics of biological samples. *Bioanalysis*, 9(1):53-65.
- Nakagawa, K., Kawana, S., Hasegawa, Y. & Yamaguchi, S. 2010. Simplified method for the chemical diagnosis of organic aciduria using GC/MS. *Journal of Chromatography B*, 878(13-14):942-948.
- Nordmann, J. & Nordmann, R. 1961. Organic acids in blood and urine. In. *Advances in clinical chemistry*. 4: Elsevier. pp. 53-120.
- Palmas, F., Mussap, M. & Fattuoni, C. 2018. Urine metabolome analysis by gas chromatography–mass spectrometry (GC–MS): Standardization and optimization of protocols for urea removal and short-term sample storage. *Clinica Chimica Acta*, 485:236-242.
- Pasikanti, K.K., Ho, P. & Chan, E. 2008. Gas chromatography/mass spectrometry in metabolic profiling of biological fluids. *Journal of Chromatography B*, 871(2):202-211.
- Plotka-Wasyłka, J., Szczepańska, N., Guardia, M.d.I. & Namieśnik, J. 2015. Miniaturized solid-phase extraction techniques. *Trends in Analytical Chemistry*, 73:19-38.
- Reinecke, C.J., Koekemoer, G., Van der Westhuizen, F.H., Louw, R., Lindeque, J.Z., Mienie, L.J. & Smuts, I. 2012. Metabolomics of urinary organic acids in respiratory chain deficiencies in children. *Metabolomics*, 8:264-283.

- Rozio, M.G., Rosato, A., Iadarola, L. & Carrara, S. 2019. Compounds Relative Response Factor, a Reliable Quantification Within Extractable Testing. *White Paper, Eurofins Biopharma Product Testing Italy*,
- Ryan, D., Robards, K., Prenzler, P.D. & Kendall, M. 2011. Recent and potential developments in the analysis of urine: a review. *Analytica chimica acta*, 684(1-2):17-29.
- Shaw, J. 2011. Fathoming Metabolism. The study of metabolites does an end run around genomics to provide telling clues to your future health. *Harv Mag*, 3:27-31.
- Shoemaker, J.D. & Elliott, W.H. 1991. Automated screening of urine samples for carbohydrates, organic and amino acids after treatment with urease. *Journal of Chromatography B: Biomedical Sciences and Applications*, 562(1-2):125-138.
- Soares da Silva Burato, J., Vargas Medina, D.A., de Toffoli, A.L., Vasconcelos Soares Maciel, E. & Mauro Lanças, F. 2020. Recent advances and trends in miniaturized sample preparation techniques. *Journal of Separation Science*, 43(1):202-225.
- Stone, J. 2017. Sample preparation techniques for mass spectrometry in the clinical laboratory. In. *Mass Spectrometry for the Clinical Laboratory*: Elsevier. pp. 37-62.
- Tsoukalas, D., Alegakis, A., Fragkiadaki, P., Papakonstantinou, E., Nikitovic, D., Karataraki, A., ... Drakoulis, N. 2017. Application of metabolomics: Focus on the quantification of organic acids in healthy adults. *International Journal of Molecular Medicine*, 40(1):112-120.
- Verhaeghe, B.J., Lefevre, M.F. & De Leenheer, A.P. 1988. Solid-phase extraction with strong anion-exchange columns for selective isolation and concentration of urinary organic acids. *Clinical chemistry*, 34(6):1077-1083.
- Wajner, M., de Moura Coelho, D., Ingrassia, R., de Oliveira, A.B., Busanello, E.N.B., Raymond, K., ... Vargas, C.R. 2009. Selective screening for organic acidemias by urine organic acid GC-MS analysis in Brazil: Fifteen-year experience. *Clinica Chimica Acta*, 400(1-2):77-81.
- Webb-Robertson, B.-J., Kim, Y.-M., Zink, E.M., Hallaian, K.A., Zhang, Q., Madupu, R., ... Metz, T.O. 2014. A statistical analysis of the effects of urease pre-treatment on the measurement of the urinary metabolome by gas chromatography-mass spectrometry. *Metabolomics*, 10:897-908.
- Wells, W., Chin, T. & Weber, B. 1964. Quantitative analysis of serum and urine sugars by gas chromatography. *Clinica Chimica Acta*, 10(4):352-359.
- Yi, L., Shi, S., Yi, Z., He, R., Lu, H. & Liang, Y. 2014. MeOx-TMS derivatization for GC-MS metabolic profiling of urine and application in the discrimination between normal C57BL/6J and type 2 diabetic KK-Ay mice. *Analytical Methods*, 6(12):4380-4387.
- Zhu, K., Gu, B., Kerry, M., Mintert, M., Luong, J. & Pursch, M. 2017. Elimination of N, O-bis (trimethylsilyl) trifluoroacetamide interference by base treatment in derivatization gas chromatography mass spectrometry determination of parts per billion of alcohols in a food additive. *Journal of Chromatography A*, 1490:74-79.

PARTS 2 AND 3: METHOD DEVELOPMENT

This section was published (Appendix A) and presented at an international conference (Appendix B):

- Olivier, C., Allen, B. and Luies, L. (2023). Optimising a urinary extraction method for non-targeted GC-MS metabolomics. *Scientific Reports*, 13(1), p.17591. (<https://doi.org/10.1038/s41598-023-44690-7>)
- Olivier, C. and Luies, L. (2023). Optimising a urinary extraction method for non-targeted GC-MS metabolomics. Presented (as poster by L. Luies) at the 19th Annual Conference of the Metabolomics Society, Niagara Falls, Canada, 18–22 June 2023.

3.6. ABSTRACT

Urine is ideal for non-targeted metabolomics, providing valuable insights into normal and pathological cellular processes. Optimal extraction is critical since non-targeted metabolomics aims to analyse various compound classes. Here, we optimised a low-volume urine preparation procedure for nontargeted GC-MS. Five extraction methods (four organic acid [OA] extraction variations and a “direct analysis” [DA] approach) were assessed based on repeatability, metabolome coverage, and metabolite recovery. The DA method exhibited superior repeatability, and achieved the highest metabolome coverage, detecting 91 unique metabolites from multiple compound classes comparatively. Conversely, OA methods may not be suitable for all non-targeted metabolomics applications due to their bias toward a specific compound class. In accordance, the OA methods demonstrated limitations, with lower compound recovery and a higher percentage of undetected compounds. The DA method was further improved by incorporating an additional drying step between two-step derivatization but did not benefit from urease sample pre-treatment. Overall, this study establishes an improved low-volume urine preparation approach for future non-targeted urine metabolomics applications using GC-MS. Our findings contribute to advancing the field of metabolomics and enable efficient, comprehensive analysis of urinary metabolites, which could facilitate more accurate disease diagnosis or biomarker discovery.

3.7. INTRODUCTION

Metabolomics refers to the quantitative measurement of dynamic metabolic changes in a system responding to genetic modifications or physiological stimuli, either external (e.g., drugs) or internal (e.g., nutrients) (Collino *et al.*, 2013). It involves the analysis of the metabolome, which can be defined as the complete set of small molecules (metabolites) present in a

biological sample. Multi-level profiling of metabolites can be conducted in an unbiased manner through the comparison of sets of biological samples (Isa *et al.*, 2018).

Gas chromatography mass spectrometry (GC-MS) is considered the “gold standard” for metabolomics studies (Fiehn, 2016) due to its many advantages, including its suitability for both volatile and non-volatile compound analyses (after derivatization), affordability in terms of low running costs, high sensitivity with excellent chromatographic and mass resolution, good dynamic range, and compound identification using mass spectral library matching (of which extensive commercial libraries are available). Furthermore, GC-MS provides additional and orthogonal data (such as retention time), reproducible chromatographic separations (making results comparable between different laboratories and analysts), and the ability to distinguish stereoisomers. GC-MS also has shorter run times with lower bleed (Christou *et al.*, 2014).

Although GC-MS has some disadvantages, such as the co-elution of compounds and slower scan rates, most of these can be compensated for, for example, with two-dimensional coupling (GCxGC) or by coupling the system to time-of-flight mass spectrometry (TOFMS), respectively (De Villiers & Du Loots, 2013). These approaches broaden the application of GC-MS substantially (Mamyrin, 2001). However, since GC-MS is limited to volatile, thermally stable compounds — and most natural occurring compounds in a biological matrix are non-volatile and unstable at high temperatures — samples need to be derivatized prior to analysis (Khodadadi & Pourfarzam, 2020). Derivatization refers to the process of chemically modifying a compound, exploiting its polarity, and boiling point, including the interaction with a GC’s capillary column to achieve compound separation. Hence, most methods require two-step derivatization, employing oximation followed by silylation (Chan *et al.*, 2011), although the latter is often used as a one-step derivatization approach. These derivatization methods will be further explored in this section.

Urine has proven to be an excellent sample matrix for non-targeted metabolomics studies because it reflects both normal and pathological cellular processes, providing a holistic assessment of metabolic profiles. Urine is a particularly valuable sample matrix due to its ease of collection, straightforward handling, and processing. These advantages make urine a promising alternative biological sample, especially given the challenges in the collection and yield of other biological samples in certain diseases (Gupta-Wright *et al.*, 2016). Early therapy intervention and improved treatment effectiveness depend on the insights derived from urine metabolomics. Therefore, urine metabolomics can be employed to rule out potential metabolic defects, assess nutritional issues, and detect signs of infections or overgrowth (including pathogen or flora). To maximise the utility of urinary metabolome data for disease characterisation and treatment strategies, effective sample preparation, coupled with selective

extractions tailored to the chosen analytical technique, becomes critically important (Madla *et al.*, 2012).

For urine GC-MS analyses, organic acid (OA) metabolite extraction remains a popular option. The principle of an OA extraction is typically based on a two-step liquid–liquid extraction (LLE; also known as partitioning) (Su *et al.*, 2017) which involves the use of two immiscible solvents (Soares da Silva Burato *et al.*, 2020). However, LLE is better suited for targeted investigations in clinical settings, such as diagnosing organic acidurias (Christou *et al.*, 2014). Optimal extraction is critical since non-targeted metabolomics aims to analyse various compound classes at varying concentrations. To this end, recent advances in technology have led to the development of more efficient and cost-effective methods for compound extraction, such as the direct analysis (DA) method, for non-targeted metabolomics.

The DA method involves deproteinisation, sample concentration, and derivatization (Esterhuizen *et al.*, 2019). Although straightforward, various studies suggest an additional drying step between two-step derivatization (Glicksberg & Kerrigan, 2020; Liebeke & Puskás, 2019) to improve reproducibility (since residual water and other contaminants can be removed), reduce matrix effects (by removing water and other impurities that can interfere with ionisation and compound detection), and increase sensitivity (since background noise is reduced while the signal-to-noise ratio is increased).

One significant challenge in non-targeted metabolomics using urine samples, is the presence of urea, which can obscure other metabolites in the sample with the same retention time and/or interfere with the derivatization process, leading to incomplete chemical transformation and the formation of urea-derived artifacts (Chan *et al.*, 2011; Khodadadi & Pourfarzam, 2020). To address this issue, various studies suggest using the enzyme urease to remove excess urea from urine samples prior to GC-MS analysis (Webb-Robertson *et al.*, 2014; Wells *et al.*, 1964). This also has the advantage of avoiding column overloading, peak distortions, lower chances of co-eluting peaks, and lower coefficient of variation (CV) values compared to urease-non-treated samples (Palmas *et al.*, 2018; Webb-Robertson *et al.*, 2014). However, the use of urease pre-treatment has its disadvantages, including the occurrence of unwanted chemical transformations and secondary enzymatic reactions that can alter the metabolic profile (Kim *et al.*, 2020; Kind *et al.*, 2007) and the inability to detect several compounds in urease-treated samples compared to non-treated samples (Webb-Robertson *et al.*, 2014). As a result, the use of urease pre-treatment in non-targeted metabolomics remains a subject of debate in the scientific community.

In this part of the investigation, we aimed to optimise a low-volume urine preparation procedure for non-targeted GC-MS analyses, covering various compound classes. In Section A Part 2, five methods were evaluated based on repeatability, metabolome coverage, and metabolite

recovery. An OA extraction method (Method 1, for which the urine volume was optimised in Section A Part 1) was compared with modified versions, namely the use of an additional extraction solvent (Method 2), two-step derivatization (Method 3), and a combination of Methods 2 and 3 (Method 4). A DA method (Method 5) was also tested. In Part 3, the superior method was further optimised by investigating if an additional drying step between two-step derivatization (Method 6) and urease sample pre-treatment (Method 7) is indeed beneficial for disease characterisation studies.

3.8. MATERIALS AND METHODS

3.8.1. Chemicals

The following reagents were acquired from Sigma Aldrich (St. Louis, Missouri, USA) and Merck (Darmstadt, Germany): 3-phenylbutyric acid, MOX-HCl, anhydrous sodium sulphate (Na_2SO_4), urease, glucose, 4-aminobutyric acid (GABA), L-alanine, L-leucine, L-phenylalanine, L-tryptophan, L-tyrosine, ascorbic acid, citric acid, succinic acid, palmitic acid, arabinose, ribose, BSTFA with 1% TMCS, and pyridine. Other chemicals used were acetonitrile, ethyl acetate, hydrogen chloride and diethyl ether from Burdick and Jackson brand (Honeywell International Inc., Muskegon, USA).

3.8.2. Standards

A spiking mixture/solution (50 ppm) was prepared prior to the experiments, using known amounts of compounds from different compound classes. This mixture was used to determine the percentage recovery of the known compounds introduced to the samples. Three different compound classes were included: amino acids (GABA, L-alanine, L-leucine, L-phenylalanine, L-tryptophan, L-tyrosine), organic acids (ascorbic acid, citric acid, succinic acid), and a fatty acid (palmitic acid). An internal standard (50 ppm) solution containing 3-phenylbutyric acid was also prepared separately.

3.8.3. Sample Preparation

A single quality control (QC) sample, compiled from ethically approved, previously collected healthy urine samples ($n=32$) from which multiple smaller (100 μL) aliquots were made, was used in this experiment. Before an aliquot was transferred, the material was thoroughly mixed by vortex. The sample aliquots were thawed at room temperature, and to eliminate any crystals and make sure that any precipitated material has dissolved, the urine was centrifuged at 15 700 $\times g$ for 5 minutes at room temperature.

3.9. METABOLITE EXTRACTION

Although a thorough overview of this study's experimental approach is provided in Chapter 2, a focused depiction of the experimental design specific to Section A's Parts 2 and 3 is available in Figure 3.6. For Part 2 of this investigation, five methods were evaluated based on repeatability, metabolome coverage, and metabolite recovery. An OA extraction method (previously optimised for a reduced urine volume; Method 1) was compared with modified versions of this method, namely the use of acetonitrile as an additional extraction solvent (Method 2), two-step derivatization to include both oximation and silylation (Method 3), as well as a combination of Methods 2 and 3 (Method 4). These methods were also compared to a DA method (Method 5).

In Part 3, the superior method was further optimised by investigating the effect(s) of an additional drying step between two-step derivatization (Method 6). Hereafter, urease sample pre-treatment was evaluated (Method 7) by comparing urease-treated samples to both urease-non-treated and water-treated samples. Furthermore, the effect of the heating process that urease pre-treatment requires was investigated by comparing the urease-treated to urease-non-treated samples subjected to both heating and sonication.

Methods 1–5 was conducted as a single experiment, for which six QC aliquots were used each (n=30), while Method 6 was conducted in a second experiment, requiring 18 aliquots, and Method 7 as a third experiment, which needed 30 aliquots.

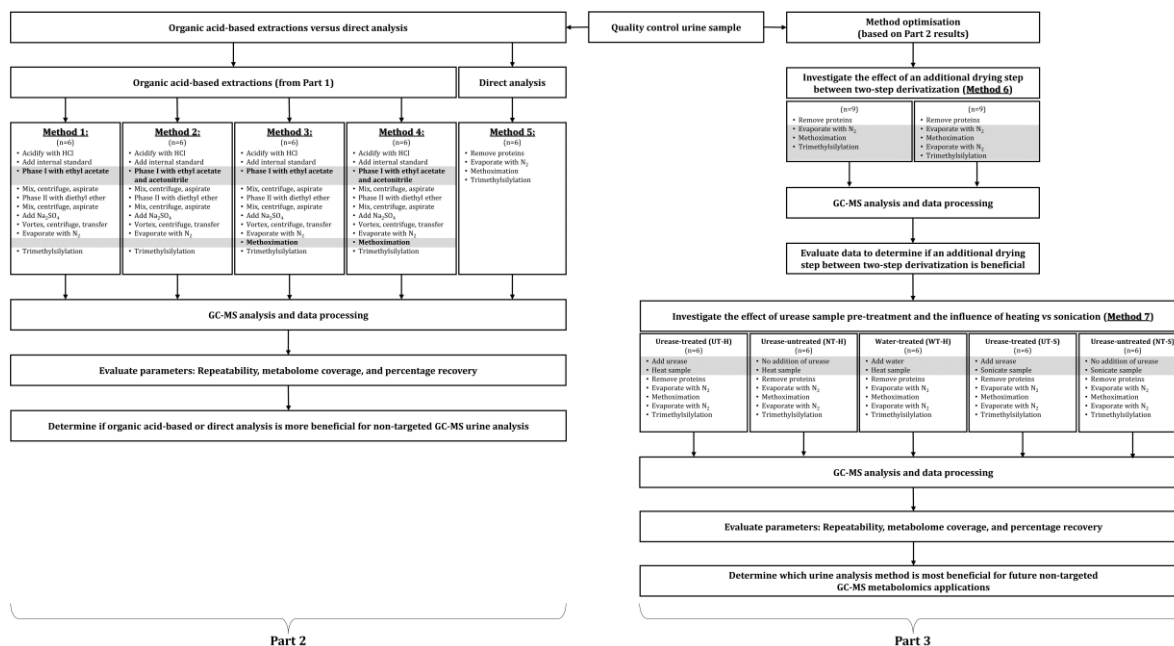


Figure 3.6: Summary of the experimental design. This study comprises a Part 2 (to evaluate five extraction methods) and Part 3 (to optimise a urine preparation procedure for non-targeted GC-MS analyses). The variations between the methods are indicated in grey sections. For Methods 1–5, six QC aliquots were used each, while Method 6 required 18 aliquots and Methods 7 needed 30 aliquots.

3.9.1. Method 1 (Traditional OA Extraction)

Six drops of 5 N HCl were added to 100 μ L urine samples filled up to 500 μ L with a 9% saline solution (n=6), followed by the addition of internal standard (100 μ L, final concentration of 50 ppm). Hereafter, half of the samples (n=3) were spiked (referred to as the “pre-spike samples”), and ethyl acetate (6 mL) was added to all samples. The samples were mixed (30 minutes) and centrifuged (845xg for 5 minutes at 4°C), and the organic phase was collected. Diethyl ether (3 mL) was added to the remaining aqueous phase, and the samples were again mixed, centrifuged, and the organic phase collected and combined with the previously collected phase. Approximately 3 g Na₂SO₄ was added to each sample’s organic phases, followed by vortex and centrifugation. The organic phase was transferred to a new tube and the remaining half of the samples were spiked (n=3; referred to as the “post-spike samples”). All samples were completely dried under a light stream of nitrogen gas at 37°C. Silylation was performed with 50 μ L BSTFA with 1% TMCS and 50 μ L pyridine, with incubation at 60°C for 60 minutes. Hereafter, samples were allowed to cool to room temperature and transferred to a GC sample vial with an insert and again capped (Luies & Loots, 2016).

3.9.2. Method 2 (extraction solvent evaluation)

This method was identical to that of Method 1, except for the addition of 1.5 mL ice-cold acetonitrile with ethyl acetate: Six drops of 5 N HCl were added to 100 μ L urine filled up to 500 μ L with a 9% saline solution (n=6). Hereafter, the internal standard (100 μ L, final concentration of 50 ppm) was added, and half of the samples (n=3) were spiked (“pre-spike samples”). Ethyl acetate (6 mL) was added to all samples, followed by ice-cold acetonitrile (1.5 mL) (Majors, 2013), whereafter samples were mixed (30 minutes) and centrifuged (845xg for 5 minutes at 4°C). After the organic phase was collected, diethyl ether (3 mL) was added to the remaining aqueous phase, and the samples were again mixed, centrifuged, and the organic phase collected and combined with the previously collected phase. Approximately 3 g Na₂SO₄ was added to the organic phases of each sample, followed by vortex and centrifugation. The organic phases were then transferred to a new tube and the remaining half of the samples were spiked (n=3; “post spike samples”). All samples were evaporated to complete dryness under a light stream of nitrogen gas at 37°C. Silylation was performed with BSTFA with 1% TMCS (50 μ L) and pyridine (50 μ L), with incubation at 60°C for 60 minutes. Hereafter, samples were allowed to cool to room temperature and transferred to a GC sample vial with an insert and again capped.

3.9.3. Method 3 (One-Step versus Two-Step Derivatization)

This method was identical to that of Method 1, except for an addition methoximation step prior to silylation (a two-step derivatization): After the addition of six drops of 5 N HCl and a 50 ppm internal standard (100 μ L) to urine samples (n=6; 100 μ L filled up to 500 μ L with a 9% saline solution), half of the samples (n=3) were spiked ("pre-spike samples"). Hereafter, and ethyl acetate (6 mL) was added to all samples, followed by mixing (30 minutes), centrifugation (845xg for 5 minutes at 4°C), and organic phase collection. Diethyl ether (3 mL) was added to the remaining aqueous phase, and the samples were again mixed, centrifuged, and the second organic phase collected and combined with the first. Na₂SO₄ (approximately 3 g) was added to the collected organic phases, followed by vortex and centrifugation. Hereafter, the samples were transferred to a new tube and the remaining half were spiked (n=3; "post-spike samples"). All samples were completely dried under a light stream of nitrogen gas at 37°C prior to methoximation. Thus, 50 μ L MOX-HCl dissolved in pyridine (20 mg/mL) was added to each sample, capped, and incubated at 50°C for 90 minutes. Hereafter, the samples were left to cool to room temperature prior to silylation with 50 μ L BSTFA with 1% TMCS (no additional pyridine was added), with incubation at 60°C for 60 minutes. The samples were again allowed to cool to room temperature and transferred to a GC sample vial with an insert and re-capped (Patti et al., 2012).

3.9.4. Method 4 (Combination of Methods 2 and 3)

This method was identical to that of Methods 2 and 3, combining the use of ice-cold acetonitrile with ethyl acetate, and two-step derivatization: To each urine sample (n=6; 100 μ L urine filled up to 500 μ L with a 9% saline solution), six drops of 5 N HCl and 100 μ L of an internal standard (500 ppm) was added. Hereafter, half of the samples were spiked (n=3; "pre-spike samples"). Ethyl acetate (6 mL) and ice-cold acetonitrile (1.5 mL) was added to all samples, followed by mixing (30 min), centrifugation (845xg for 5 minutes at 4°C), and organic phase collection. Diethyl ether (3 mL) was then added to the remaining aqueous phase, again followed by mixing, centrifugation, and the organic phase collection (which was combined with the previously collected phase). Na₂SO₄ (approximately 3 g) was added to the collected organic phases, and after vortex and centrifugation, the samples were transferred to a new tube. The remaining half of the samples were then spiked (n=3; "post-spike samples"). All samples were dried with nitrogen gas at 37°C. Once samples were completely dry, methoximation was performed using 50 μ L MOX-HCl dissolved in pyridine (20 mg/mL). Samples were capped and incubated at 50°C for 90 minutes. Hereafter, the samples were left to cool to room temperature prior to silylation with 50 μ L BSTFA with 1% TMCS (no additional pyridine was added), with incubation at 60°C for 60 minutes. The samples were again allowed to cool to room temperature, before transferring them to a GC sample vial with an insert, and re-capped.

3.9.5. Method 5 (DA)

The internal standard (100 μ L; final concentration of 50 ppm) was added to 100 μ L urine. Half of the samples were spiked ($n=3$; “pre-spike samples”), ice-cold acetonitrile (300 μ L) was added, and samples were centrifuged at room temperature (15000 $\times$ g for 5 minutes). The supernatant was collected and transferred to a GC-MS vial, whereafter the remaining half of the samples were spiked ($n=3$; “post-spike samples”). Similar as before, the samples were dried completely under a light stream of nitrogen gas at 37°C. The vials were allowed to cool to room temperature before applying two-step derivatization: Samples were methoxylated using 50 μ L MOX-HCl dissolved in pyridine (20 mg/mL), capped, and incubated at 50°C for 90 minutes. After cooling to room temperature, samples were silylated with 50 μ L BSTFA with 1% TMCS (no additional pyridine was added), capped, and incubated at 60°C for 60 minutes. Samples were transferred to a glass insert, placed into the original GC-MS vial, and capped (Esterhuizen et al., 2019).

3.9.6. Method 6 (Additional Drying Step)

For this part of the investigation, 18 QC aliquots were used, divided into two groups: Group 1 samples ($n=9$) were subjected to Method 5 exactly as indicated above (including two-step derivatization). Group 2 samples ($n=9$) were also subjected to Method 5; however, the derivatization process was adjusted by adding another drying step prior to silylation. Following methoximation as described above (50 μ L MOX-HCl dissolved in pyridine [20 mg/mL]), samples were again dried under a light stream of nitrogen gas at 37°C prior to silylation with 50 μ L BSTFA with 1% TMCS (no additional pyridine was added). After cooling to room temperature, samples were transferred to a glass insert, placed into the original GC-MS vial, and capped (Liebeke & Puskás, 2019).

3.9.7. Method 7 (Urease Pre-Treatment)

QC aliquots ($n=30$) were divided into five groups, namely Group 1 urease-treated samples subjected to heating ($n=6$; UT-H); Group 2 urease-non-treated samples subjected to heating ($n=6$; NT-H); Group 3 water-treated samples subjected to heating to serve as a blank sample for investigating the diluting effect urease pre-treatment has on samples ($n=6$; WT-H); Group 4 urease-treated samples subjected to sonication instead of heating ($n=6$; UT-S); Group 5 urease-non-treated samples subjected to sonication ($n=6$; NT-S). The analysis of these samples was based on Method 6, however, following the addition of the internal standard, urease (100 μ L of a 1 mg/mL solution; Groups 1 and 4) or water (100 μ L; Group 3) was also added. The samples were then either heated (30 minutes at 37°C; Groups 1–3) (Webb-Robertson et al., 2014) or sonicated (30 minutes at room temperature with 005 power [150 W]

using a model 703 sonicator with 5.7 L capacity; Groups 4 and 5) (Palmas *et al.*, 2018) for adequate reaction to occur, before continuing with the addition of ice-cold acetonitrile whereafter the rest of the method proceeded as described previously (see Section 3.9.6 “Method 6”). In this study, a sonication frequency of 50 Hz was used to facilitate enzyme action without significantly disrupting cellular integrity (Rokhina *et al.*, 2009). Despite potential heat generation from mechanical vibrations, the temperature during sonication was not monitored, as the chosen conditions are based on protocols demonstrated to prevent significant thermal degradation of sensitive compounds (Palmas *et al.*, 2018).

3.10. GC-MS ANALYSIS

Prior to sample analysis, a routine clean-up and maintenance check was performed (i.e., leak check, tune check, and mass calibration) and a new liner and septum was inserted to prevent any undesired reactions and surface adsorption phenomena. Samples were analysed in two batches (Part 2; Methods 1–5) and (Part 3; Methods 6 and 7), in a randomised manner. The samples were placed into an auto-sampler tray (Gerstel MS Germany), combined with a Pegasus 4D GCxGC-TOFMS system (LECO Africa [Pty] Ltd, Johannesburg, South Africa) and fitted with an Agilent 7890 GC and TOFMS (LECO Africa). All samples (1 μ L) were injected using a 1:10 split ratio with purified helium as a carrier gas, set at a constant flow of 1.4 mL/min. Chromatographic separation was achieved with a Restek Rxi-5MS primary capillary column (28.2 m; 250 μ m diameter; 0.25 μ m film thickness), and a Restek Rxi-17 secondary capillary column (1.3 m; 250 μ m diameter; 0.25 μ m film thickness). Throughout the entire chromatographic run, the front inlet temperature was maintained at 250°C, the transfer line at 225°C, and the ion source was kept constant at 200°C. Cryomodulation was achieved with a hot pulse of nitrogen gas for 0.7 s, every 3 s. The primary oven programme (see Figure S1 in the supplementary information of Appendix A) started at 70°C, which was held for 1 minute. The temperature was then ramped up as follows: 5°C/min to 100°C, 10°C/min to 160°C, 13°C/min to 230°C, and finally 20°C/min to 300°C, which was held for 2 minutes. The secondary oven was programmed identical to that of the primary oven, except for a +5°C at each interval. The total run time of each sample was $\approx$ 24 minutes. Prior to MS data acquisition, a 480 s solvent delay was used, during which no mass spectra were captured. However, to accurately represent retention times, this interval was appended to the time axis of the GC column. Mass spectra were collected over a range of 50–800 m/z at an acquisition rate of 20 spectra per second. The filament bias was -70 eV, and the detector voltage was optimised at an offset of 50 V.

3.11. DATA ANALYSIS

For each experimental dataset, Leco ChromaTOF-GC Software (v4.72.0.0) was used for MS deconvolution (with a signal-to-noise ratio of 100 and a minimum of three apexing peaks), peak identification (with a 60% similarity match) and peak alignment. The raw data from the GCxGC-TOFMS was exported as a .csv file and converted to an Excel Workbook (.xlsx) format. The data was normalised based on the internal standard, followed by normalisation based on the creatinine value. Various parameters were assessed using Excel, including analytical repeatability, metabolome coverage, and metabolite recovery.

3.12. RESULTS AND DISCUSSION

In Part A of this investigation, five methods were evaluated based on repeatability, metabolome coverage, and metabolite recovery. The goal here was to select the most appropriate method for non-targeted GC-MS application. Hereafter, the selected method was further explored for possible optimisation in Part 3.

3.12.1. Repeatability

To determine the analytical repeatability for Methods 1–5, the first three samples of each method (i.e., the “pre-spike samples”) were used to calculate the CV values, expressed as percentage. When interpreting the CV values, the more overall compounds with a CV < 50%, the better the method (Schoeman *et al.*, 2012). Figure 3.7(a) illustrates the CV distribution for all detected peaks, for the five extraction methods. When considering the analytical repeatability, Method 5 had the best repeatability (82.4% of all detected peaks had a CV < 50%), followed by Method 3 (62.1%), Method 4 (61.9%), Method 1 (54.6%), and lastly Method 2 (36.4%).

For further confirmation of the above ranking, the web-based server MetaboAnalyst was used to statistically evaluate the data and compile a principal component analysis (PCA) scores plot [see Figure 3.7(b)]. This plot provides an overview of the overall similarity or dissimilarity among all five extraction methods without directly incorporating the CV values, but rather considering patterns and relationships among the samples of each method based on their measured variables. Indeed, the PCA analysis revealed distinct clustering patterns among the five methods. Specifically, the OA extraction methods (Methods 1–4) exhibited a tight clustering, suggesting similar metabolite profiles. The DA samples (Method 5) were positioned distinctly and significantly distant from the OA extraction methods. This observation suggests substantial dissimilarity in the metabolite profiles of Method 5 compared to the other methods, likely due to its unique extraction principle and variability in its metabolite composition. On a

PCA, the approach with the least amount of variance will have the smallest 95% confidence region circle, which is where the samples cluster closest together. Thus, the bigger the 95% confidence region, the less repeatable the method. The large green ellipse, which represents Method 2 (i.e., the addition of acetonitrile), shows great variance between the samples. Although acetonitrile has a superior extraction efficiency (Majors, 2013), also observed that its extracts were still somewhat “dirty”, possibly contributing to the observed variation. Furthermore, the presence of water in acetonitrile is a potential factor in the variance between the extracted analytes, as it may impact how effectively different OA's are extracted, depending on how the OA is distributed between the aqueous and organic phases (Phiri, 2017). This may lead to the loss of compounds during extraction, making this method the least repeatable. Thus, it may be beneficial to add kosmotropic salts (e.g., MgSO_4 or NaCl) to phase separate acetonitrile from a water/acetonitrile solution. The original in-house method samples (Method 1; red ellipse), has slight variation between the samples but less so compared to Method 2. Considering the additional compound stabilisation that oximation offers (Patti et al., 2012), Method 3 (dark blue ellipse) and Method 4 (blue ellipse) have better confidence region circles than Methods 1 and 2, indicating less variability, which concurs with Figure 3.7(a). Similarly, Method 5 (pink ellipse), which also entails oximation prior to silylation, shows tight sample clustering. Thus, the method ranking based on the PCA scores plot corresponds with that of the CV-graph.

Based on Figure 3.7, Method 5 yields more repeatable results compared to any of the OA extraction methods. This is possibly due to oximation for compound stabilisation, in conjunction with its easy preparation (i.e., reduced analytical steps involved) since the OA extraction methods are more labour-intensive and hence more prone to errors.

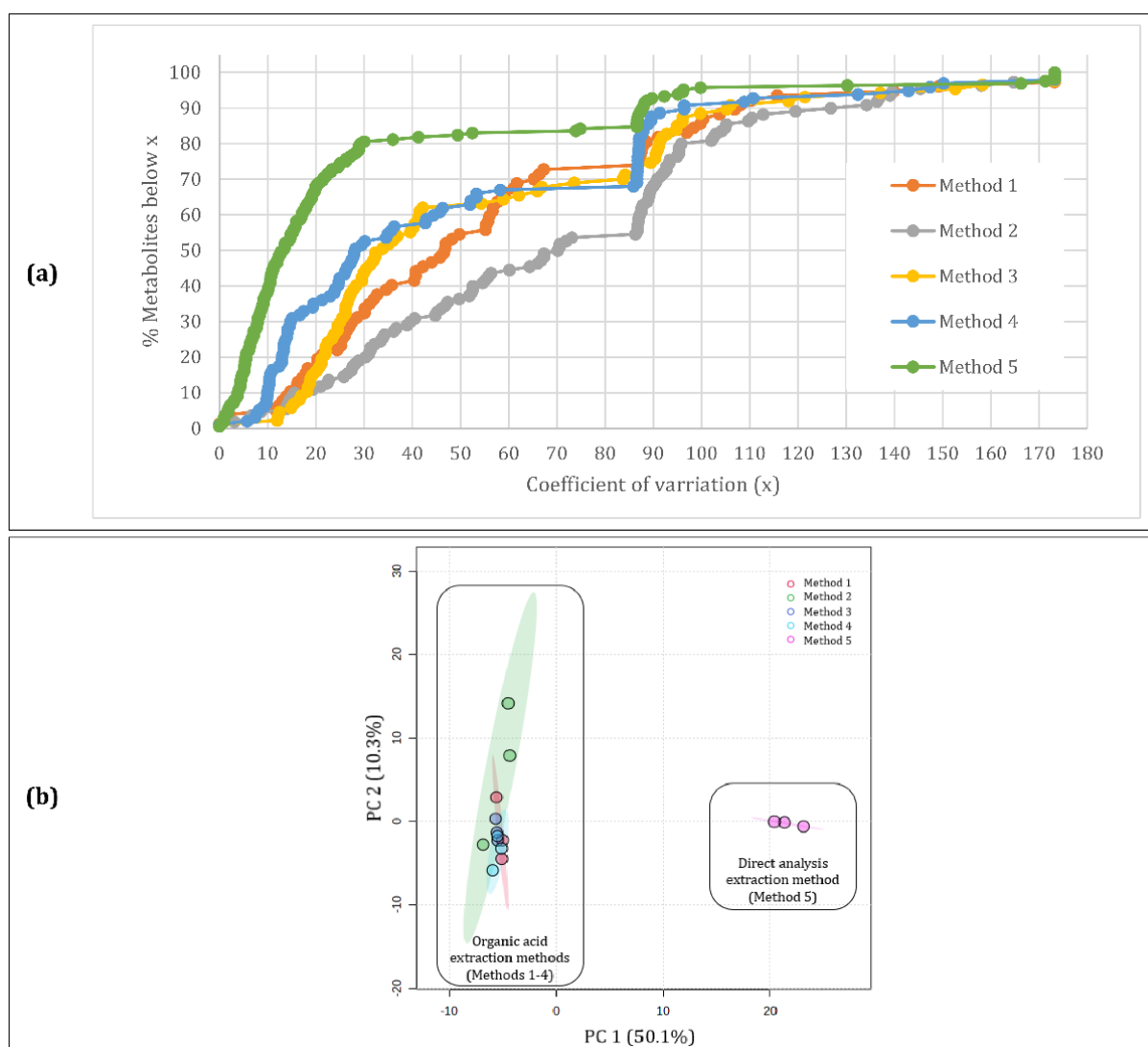


Figure 3.7: Evaluation of repeatability of the five extraction methods. (a) The percentage of metabolites below a coefficient of variation (CV) threshold < 50%, ranked as Method 5 (82.4%), Method 3 (62.1%), Method 4 (61.9%), Method 1 (54.6%), and Method 2 (36.4%). (b) Principal component analysis (PCA) scores plot which illustrates the clustering patterns among the samples based on their metabolite profiles/composition. The plot provides an overview of the overall similarity or dissimilarity among the methods' samples without directly incorporating the CV values.

3.12.2. Metabolome Coverage

The metabolome coverage of each method was assessed using the normalised “pre-spike samples” only. The total number of peaks detected across all five extraction methods (n=220) was determined, whereafter the number of overall peaks not detected in at least 50% of all samples for each extraction method was calculated as: % peaks undetected = $\frac{\text{Number of peaks with zero average}}{220 \text{ peaks overall}} \times 100$. The extraction methods were then ranked based on this calculated percentage, from the least number of undetected peaks to the most undetected peaks as follows: Method 5 (26.4% undetected); Method 2 (51.4% undetected); Method 4

(57.3% undetected); Method 3 (61.8% undetected); and Method 1 (66.4% undetected). Upon further investigation, we found that 91 peaks were exclusively detected using Method 5, whereas 58 were specific to the four OA extraction methods. Of these 58, 13 were exclusive to Method 2, seven were exclusive to Method 4, one was exclusive to Method 3, and none was exclusive to Method 1 (see Table 3.4 for complete lists of these unique compounds). As can be seen, both metabolome coverage rankings (either considering the percentage of undetected peaks or the number of unique peaks) concurred.

3.12.3. Metabolite Recovery

A spiking mixture is a carefully prepared solution containing known concentrations of specific compounds (see Section 3.8.2 “Standards”), and serves as a vital component in analytical chemistry, fulfilling various purposes. In this study, a spiking mixture was used to specifically assess the recovery rate of target analytes during sample preparation and analysis. This assessment involves adding a known amount of analyte into the natural test sample matrix. By comparing the amount of spike compounds added to the sample with the amount detected in the final analysis, researchers can calculate the recovery percentage, which indicates the efficiency of target analyte extraction and measurement. Thus, metabolite recovery was determined by comparing the concentrations of compounds in the “pre-spike” and “post-spike” samples, using only those compounds included in the spiking solution (see Section 3.8.2 “Standards”). The percentage of recovery was calculated as: % recovery =

$$\frac{\text{Pre-spike concentration}}{\text{Post-spike concentration}} \times 100 \text{ (see Table 3.5).}$$

Table 3.4: Unique compounds detected for each method, and those detected in only the organic acid extraction methods.

Method name and number of unique peaks	List of unique peaks/compounds			
Method 1 (n=0)	None			
Method 2 (n=13)	Analyte 003 Analyte 013 Analyte 121 Analyte 128	Analyte 135 Analyte 148 Analyte 150	Analyte 156 Analyte 157 Analyte 167	Analyte 171 Analyte 192 Analyte 308
Method 3 (n=1)	Analyte 322			
Method 4 (n=7)	2,3,4-Trihydroxy-butyraldehyde O-methyl-oxime	Analyte 036 Analyte 071	Analyte 209 Analyte 216	Analyte 325 Malonic acid
Method 5 (n=91)	2-Amino-benzoic acid 3,4,5-Trihydroxy-pentanoic acid 3,4-Dihydroxy-butyric acid 3R,4S-Tetrahydrofuran-2,3,4-triol 4-Hydroxy-pyrrolidine-2-carboxylic acid 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [1] 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [2] 5-Oxo-pyrrolidine-2-carboxylic acid Analyte 004 Analyte 005 Analyte 023 Analyte 025 Analyte 029 Analyte 035 Analyte 038 Analyte 043 Analyte 046	Analyte 050 Analyte 057 Analyte 058 Analyte 061 Analyte 063 Analyte 064 Analyte 070 Analyte 072 Analyte 085 Analyte 086 Analyte 087 Analyte 090 Analyte 096 Analyte 099 Analyte 118 Analyte 125 Analyte 126 Analyte 131 Analyte 137 Analyte 145 Analyte 151 Analyte 177 Analyte 183 Analyte 186 Analyte 199	Analyte 210 Analyte 212 Analyte 218 Analyte 220 Analyte 225 Analyte 251 Analyte 253 Analyte 266 Analyte 276 Analyte 278 Analyte 280 Analyte 285 Analyte 287 Analyte 302 Analyte 305 Analyte 306 Analyte 311 Analyte 321 Analyte 323 Arabinofuranose Arabinose Ascorbic acid Aucubin Butane-2,3-diol Creatinine	Dehydroabietic acid Erythronic acid γ -lactone Galactose [1] Galactose [2] Glucose Glucuronic acid Glucuronic acid γ -lactone Glycerol Isocitric lactone meso-Erythritol Methyl galactoside Methylmalonic acid Myo-Inositol [1] Myo-Inositol [2] N-Acetyltyrosine N-benzoyl-glycine Oxalic acid Oxalic acid [1] Oxalic acid [2] Rhamnose Ribofuranose Tagatofuranose Threo-3-deoxy-pentonic acid Xylose
Peaks unique to the organic acid methods (Methods 1–4)* (n=58)	2,3,4-Trihydroxy-butyraldehyde O-methyl-oxime 2,4-Dihydroxybutyric acid 2-Hydroxy-pentanedioic acid 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [1] 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [2] Analyte 003 Analyte 013 Analyte 014 Analyte 020 Analyte 022 Analyte 024	Analyte 027 Analyte 031 Analyte 036 Analyte 049 Analyte 052 Analyte 055 Analyte 059 Analyte 071 Analyte 121 Analyte 123 Analyte 128 Analyte 135 Analyte 143 Analyte 144 Analyte 148 Analyte 150	Analyte 156 Analyte 157 Analyte 167 Analyte 168 Analyte 170 Analyte 171 Analyte 173 Analyte 174 Analyte 175 Analyte 176 Analyte 192 Analyte 195 Analyte 205 Analyte 209 Analyte 216 Analyte 226	Analyte 239 Analyte 240 Analyte 255 Analyte 297 Analyte 301 Analyte 308 Analyte 314 Analyte 322 Analyte 325 Benzene-1,2,3-triol Citric acid [miss ID] Malonic acid diethyl ester Octadecanoic acid Propane-1,2,3-triol [1] Tetrahydro-pyran-2,3,4,5-tetraol (isomer 2)

*Some of these compounds were detected across multiple organic acid extraction methods, hence they are not unique to a specific organic acid method (Methods 1–4), however, they were not detected using the direct analysis method (Method 5).

Table 3.5: Metabolite recovery of the compounds in the spiking mixture (containing pure, known standard concentrations of different class compounds) indicated for each method. The compounds are grouped per class.

Spiking compound	Method 1 (in-house)			Method 2 (acetonitrile addition)			Method 3 (two-step derivatization)			Method 4 (combined methods 2 & 3)			Method 5 (direct analysis)		
	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery
4-Aminobutyric acid (GABA)	0	0.74	0	0	0.95	0	0	0.84	0	0	0.23	0	0.11	0.35	30.46
L-Alanine	0	3.06	0	0	4.01	0	0	3.7	0	0	1.79	0	1.85	3.01	61.35
L-Leucine	0	1.54	0	0	0.18	0	0	1.47	0	0	0.23	0	0.45	0.99	45.66
L-Tryptophan	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0
L-Tyrosine	0	0.28	0	0	0.24	0	0	0.37	0	0	0.17	0	0	0.03	0
L-Phenylalanine	0	0.53	0	0	0.65	0.39	0.09	0.78	11.94	0.03	0.45	6.15	0.21	0.24	86.55
Ascorbic acid	0	0.02	0	0	0.04	0	0	0.09	0	0	0.01	0	0.09	0.11	86.49
Citric acid	0.57	0.79	72.25	1.02	1.24	82.56	0.59	0.89	65.61	0.9	0.69	130.81	1.63	1.38	118.12
Succinic acid	0.11	0.13	82.33	0.13	0.15	85.62	0.11	0.14	80.08	0.1	0.12	79.05	0.16	0.21	78.77
Palmitic acid	0.14	0.21	67.57	0.38	0.21	181	0.17	0.34	50.67	0.12	0.18	67.16	0.09	0.13	69.28
Average recovery			22.22			34.96			20.83			28.32			57.67

Among the evaluated methods, Method 5 exhibited an average percentage recovery of 57.7%. Notably, citric acid displayed the highest recovery (118.1%), followed by phenylalanine (86.6%), ascorbic acid (86.5%), succinic acid (78.8%), palmitic acid (69.3%), alanine (61.4%), leucine (45.7%), and GABA (30.5%). For Method 2, the average recovery percentage was 35%, with the highest recovery observed for palmitic acid (181%), followed by succinic acid (85.6%), citric acid (82.6%), and phenylalanine (0.4%). Method 4 achieved an average recovery percentage of 28.3%, with citric acid showing the highest recovery (130.8%), followed by succinic acid (79.1%), palmitic acid (67.2%), and phenylalanine (6.2%). In the case of Method 1, the average percentage recovery was 22.2%, with succinic acid demonstrating the highest recovery (82.3%), followed by citric acid (72.3%) and palmitic acid (67.6%). Method 3 exhibited the lowest average percentage recovery at 20.8%, with succinic acid achieving the best recovery (80.1%), followed by citric acid (65.6%), palmitic acid (50.7%), and phenylalanine (27%). It is noteworthy that some compounds showed recoveries above 100%. This can possibly be attributed to the fact that the QC samples used during analysis were compiled from human urine rather than synthetic urine, ensuring a realistic representation in terms of robustness during this investigation. Thus, many of these compounds may already be present in the samples. Another reason, however, might be that losses occurred in the post-spiked samples with less of the spiked compounds being derivatized compared to the pre-spiked samples.

Alanine, GABA, ascorbic acid, and leucine exhibited poor extraction efficacy across all methods, resulting in a 0% recovery, except for Method 5 (DA), which extracted all these metabolites. It is well-known that the polarity of OA's can significantly influence their extraction efficiency. Acidic compounds containing more hydroxyl groups tend to have lower recovery efficacy in apolar solvents, as observed with ascorbic acid, which is very acidic with four hydroxyl groups. Mouskeftara *et al.* (2021) also reported significantly lower GABA signals when using ethyl acetate for analyte extraction compared to an acidic extraction with methyl tert-butyl ether (*MTBE*) or a more commonly used approach involving methanol (*MeOH*). Their study suggested that a combination of *MeOH* and methoxyamine with *N*-methyl-*N*-(trimethylsilyl)-trifluoroacetamide (*MSTFA*) 1% *TMCS* derivatization generally enhanced efficiency. However, *BSTFA* 1% *TMCS* exhibited a preference for certain acids, while *MSTFA* 1% *TMCS* yielded lower peaks for succinic acid. In contrast, our study achieved a recovery efficiency of 79% or more across all the OA extraction methods using *BSTFA* 1% *TMCS*. For future studies employing an *LLE* approach, exploring the potential use of *MTBE* in combination with acetonitrile is worth considering. Additionally, Pasikanti *et al.* (2008) compared various derivatization agents [*BSTFA*, *MSTFA*, and *N*-methyl-bis (trifluoroacetamide) (*MBTFA*)] used in urinary GC-MS metabolomics. In terms of the number of observed peaks, peak intensity, and reproducibility, both *BSTFA* and *MSTFA* demonstrated equal derivatization efficiency.

Since MBTFA is more effective in derivatizing secondary and tertiary amine groups, they also explored the combination use of MBTFA with BSTFA. When MBTFA derivatization was preceded by BSTFA derivatization, more amino acids were detected, however, the integrated peak areas had poor reproducibility. Based on these findings, the authors recommended a preference for BSTFA as the derivatization agent.

Tryptophan and tyrosine had 0% recovery across all extraction methods, however, this was expected since GC systems are not ideal for detecting amino acids with great precision (Blau *et al.*, 2014; Phiri, 2017). When comparing the various OA extraction techniques, Methods 3 and 4, which included an added oximation step, yielded the highest recovery of phenylalanine. This suggests that the oximation process plays a crucial role in the extraction efficiency of phenylalanine. Although oximation is primarily applied to stabilise α -keto acids and aldehydes by converting them into oximes, its effect on phenylalanine can be attributed to the stabilisation of any keto acid derivatives formed during the extraction process. Phenylalanine, through its metabolic pathways, can be transaminated to form phenylpyruvate, a keto acid. The oximation of this keto acid can lead to more stable derivatives, enhancing the extraction and detection of phenylalanine or its derivatives in GC systems (Fiehn *et al.*, 2000; Kanani *et al.*, 2008; Mouskeftara *et al.*, 2021). The metabolite recovery ranking, however, shows that acetonitrile is required for higher extraction efficacy (Majors, 2013), and that oximation alone is insufficient to recover broad ranges of metabolite classes. This highlights the complexity of metabolite extraction and the need for a combination of techniques to optimise recovery for different metabolite classes.

However, it is important to exercise caution when interpreting metabolomics results, with particular attention to the specific methodology employed for sample extraction, as well as the analytical platform used for analysis. In the field of metabolomics, the choice of analytical platform is a critical consideration when designing a study. Each platform possesses its unique strengths and limitations in terms of detecting specific metabolites or metabolite classes. For instance, nuclear magnetic resonance (NMR) is particularly adept at detecting sugars, while liquid chromatography (LC) platforms excel in the analysis of amino acids, which GC systems may have limitations in detecting with high precision. However, GC platforms offer high efficiency, sensitivity, and reproducibility, making them well-suited for conducting non-targeted or so-called “discovery mode” metabolomics research where a broad range of metabolites needs to be analysed. In the context of this study, our focus was primarily on the application of methods optimised for a GC-MS platform, which has its own unique advantages. Researchers should carefully select the analytical platform that aligns with the specific metabolites of interest and the objectives of their metabolomics study to ensure meaningful and accurate results. While our methods may not be universally applicable, they provide valuable insights into the efficient extraction and analysis of metabolites using a GC-MS

platform, and future studies exploring other platforms for specific metabolites or metabolite classes are certainly warranted.

3.12.4. Selecting The Most Appropriate Method for Non-Targeted GC-MS Application

After each validation parameter was investigated, each method was ranked based on its overall performance, and a cumulative score was calculated for each method (see Table 3.6). Since a top ranking would result in a lower cumulative score, the method with the lowest cumulative score should theoretically be the superior method.

Table 3.6: Method rankings based on validation parameters, and an overall cumulative score of each.

	<i>Repeatability</i>	<i>Metabolome coverage</i>	<i>Metabolite recovery</i>	<i>Cumulative score</i>
Method 1 (original in-house)	4	5	4	13
Method 2 (acetonitrile addition)	5	2	2	9
Method 3 (two-step derivatization)	2	4	5	11
Method 4 (combination of Methods 2 & 3)	3	3	3	9
Method 5 (direct analysis)	1	1	1	3

The ranking provided here can be considered for future studies related to semi-targeted metabolomics (Dunn, 2008). It should be noted that repeatability is subjective to the analyst's ability. Naturally, the fewer analytical steps involved, the simpler the method would be, and hence perform better in terms of repeatability. When an investigation requires a semi-targeted approach to extract OAs, for example, Methods 2 or 4 should be considered for optimal results. Furthermore, both options allowed for a high number of unique compounds and the overall recovery of organic- and fatty acids was good for both methods. On the other hand, a non-targeted strategy examines the complete metabolic state (Shulaev, 2006) and would therefore potentially benefit more from taking the percentage of undetected compounds into account, as the investigation would need to identify multiple groups of substances or different compound classes simultaneously (Nicholson *et al.*, 2002).

Considering the aim of the investigation, i.e., to optimise a low-volume urine preparation procedure for non-targeted GC-MS analyses covering various compound classes, Method 5 outperformed all other methods, consistently ranking first in every validation parameter. Therefore, we attempted to further optimise Method 5 by determining if an additional drying step between two-step derivatization (Method 6), and urease sample pretreatment, also considering the effect of heating versus sonication (Method 7), are beneficial.

3.12.5. Additional Drying Between Two-Step Derivatization

For this part of the investigation, samples were extracted as described in Section 3.9.6 (“Method 6”). When considering the total ion chromatograms (Figure 3.8) from Sample A (orange; traditional two-step derivatization) and Sample B (green; additional drying between two-step derivatization), the additional drying step seems to reduce the signal of urea significantly, which is considered advantageous.

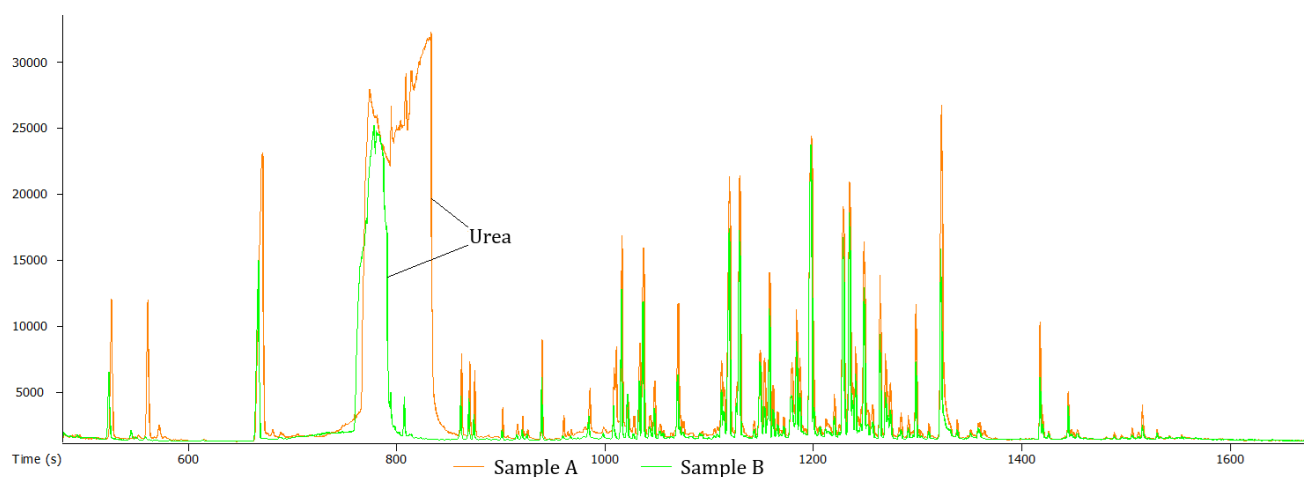


Figure 3.8: The total ion chromatograms showing the effect of an additional drying step. Sample A (orange) was subjected to two-step derivatization while Sample B (green) was subjected to an additional drying step between two-step derivatization. Sample B showed a significant reduction in the signal of urea, which is considered advantageous.

Although the average peak height was visually higher in the conventional two-step derivatization samples, certain metabolite classes still displayed higher reported areas. One such group was amino acids, which has also been reported by Liebeke and Puskás (2019) to be higher following an additional drying step. Moreover, the inclusion of an additional drying step results in a reduction of signal for specific metabolites often referred to as “junk compounds”, which commonly arise from reagent artifacts.

3.12.6. Urease Sample Pre-Treatment

To evaluate the potential advantages of urease sample pre-treatment and examine the impact of heating versus sonication, sample analysis was conducted following the procedure outlined in Section 3.9.7 (“Method 7”). The five previously mentioned sample groups (UT-H, NT-H, WT-H, UT-S, and NT-S) were compared in two separate experimental sets. The previously described criteria of repeatability (Section 3.12.1, metabolome coverage (Section 3.12.2), and metabolite recovery (Section 3.12.3) were employed to assess the outcomes of both experiments. Firstly, the effect of urease pre-treatment (during which samples were subjected

to heating only, i.e., UT-H, NT-H, and WT-H) was evaluated. The results obtained in every validation parameter is indicated in Table 3.7. Here, the water-treated samples had the lowest cumulative ranking score between the three groups, and hence performed the best, with the urease-treated samples having the most unfavourable ranking in each assessment.

Table 3.7: Methods rankings based on validation parameters, when comparing urease-treated, urease-nontreated, and water-treated samples.

	<i>Repeatability</i>	<i>Metabolome coverage</i>	<i>Metabolite recovery</i>	<i>Cumulative score</i>
Urease-treated samples (UT-H)	3	3	3	9
Urease-non-treated samples (NT-H)	2	2	2	6
Water-treated samples (WT-H)	1	1	1	3

Urease treatment effectively reduces the urea content in urine samples, but it can have additional effects beyond urea reduction, which can negatively impact sample performance. These effects are influenced by various factors. Firstly, undesired enzymatic reactions may occur since enzymes can catalyse reactions beyond the intended target, leading to chemical changes in the sample. These changes may include alterations in the metabolite profiles, degradation of certain compounds, or the generation of new metabolites as by-products of the enzymatic reactions. Such effects can impact the composition, stability, and integrity of the metabolites of interest, introducing variability and decreasing the accuracy of the analysis (Kim *et al.*, 2020).

Secondly, urease treatment may decrease the recovery of metabolites of interest due to chemical transformations or interactions with the enzymatic products generated during urea hydrolysis. If certain metabolites are poorly recovered or lost entirely due to urease treatment, it can result in gaps or biases in the metabolome coverage, compromising the accuracy and reliability of the analysis. Consequently, a comprehensive understanding of the metabolic profile may be hindered, potentially leading to incomplete or misleading interpretations of the results (Khodadadi & Pourfarzam, 2020; Webb-Robertson *et al.*, 2014).

Thirdly, water-treated samples may exhibit a slight dilution effect, enhancing the extraction capability and improving the detection and measurement of metabolites (Webb-Robertson *et al.*, 2014). Dilution helps mitigate potential matrix effects or sample interferences that could affect the analysis. Matrix effects can occur when the components of the sample, such as salts or proteins, interfere with the detection and measurement of analytes. By diluting the sample, the concentrations of these interfering components are reduced, potentially minimizing their impact on the analysis (Kruve *et al.*, 2009). Consequently, a more accurate and reliable measurement of the target analytes can be achieved. However, the optimal concentration

range depends on the specific technique and analyte characteristics, as extremely high or low concentrations may present challenges in terms of sensitivity, linearity, or dynamic range.

Fourthly, urease activity itself can cause interference or unwanted changes in the sample due to its inherent catalytic activity (Mobley, 2001; Qin & Cabral, 2002). This interference can manifest as chemical interactions or alterations in the stability, integrity, or concentration of metabolites of interest.

Lastly, urease treatment adds complexity and additional steps to the sample preparation process, increasing the risk of errors or variability. Each additional step presents an opportunity for introducing artifacts, contamination, or analyte loss (Webb-Robertson *et al.*, 2014). Inaccurate or inconsistent handling, mixing, or transferring of samples during urease pre-treatment can impact the accuracy, precision, and reproducibility of the analysis.

The total ion chromatograms (Figure 3.9) of a urease-treated sample (Sample A; green) and a urease-non-treated sample (Sample B; orange) were compared. Based on visual inspection, urease pre-treatment indeed reduced the amount of urea significantly. However, urease also reduced the overall metabolic signals, likely due to the various factors discussed above. This again highlights the detrimental effect that urease pre-treatment can have when non-targeted metabolomics studies are performed. In disease characterisation studies, for example, the metabolic profiles of patients need to remain unchanged, hence urease pre-treatment would not be recommended for such patient samples.

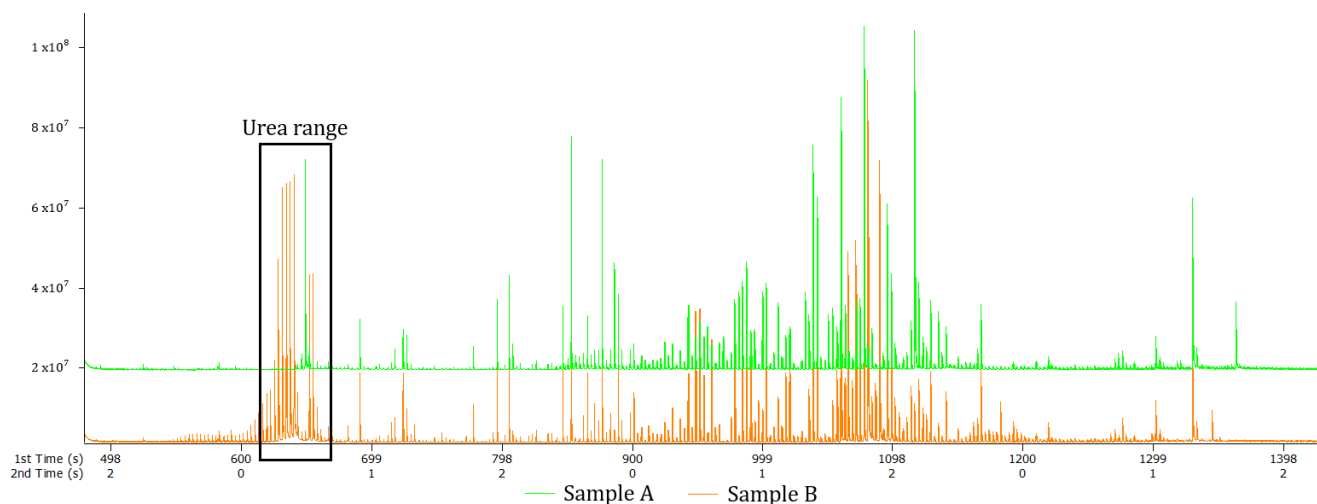


Figure 3.9: The total ion chromatograms showing the effect of urease sample pre-treatment. Sample A (green) was subjected to urease pre-treatment while Sample B (orange) shows a urease-non-treated sample. A baseline offset of 2×10^7 was used for visualisation.

Hereafter, an additional experiment was conducted to examine the impact of heating versus sonication during urease pre-treatment to determine if this could increase the benefits of such

an approach. Thus, UT-H, NT-H, UT-S, and NT-S were compared (see Section 3.9.7 “Method 7”) and again assessed based on the previously described criteria of repeatability (Section 3.12.1), metabolome coverage (Section 3.12.2), and metabolite recovery (Section 3.12.3), results of which are indicated in Table 3.8.

Table 3.8: Method rankings based on validation parameters, when comparing heating and sonication following urease pre-treatment.

	<i>Repeatability</i>	<i>Metabolome coverage</i>	<i>Metabolite recovery</i>	<i>Cumulative score</i>
Urease-treated: Heating (UT-H)	4	4	4	12
Urease-treated: Sonicate (UT-S)	2	3	2	7
Urease-non-treated: Heating (NT-H)	1	2	3	6
Urease-non-treated: Sonicate (NT-S)	3	1	1	5

Based on the results, an approach of urease non-treatment, regardless of heating or sonication, consistently outperforms approaches where samples are subjected to urease pre-treatment, demonstrating superior performance, and yielding more reliable results. These findings align with the conclusions reached by Kim *et al.* (2020), who observed that urease pre-treatment may not be as advantageous as initially believed. Their study revealed that urease pre-treatment introduced artefacts into the metabolite profiles, potentially leading to misinterpretation of results. Other studies, however, observed that urease pre-treatment samples exhibited higher concentrations compared to urease non-treated (Palmas *et al.*, 2018) and water-treated samples (Webb-Robertson *et al.*, 2014). Nonetheless, Palmas *et al.* (2018) observed a small number of metabolites in urease non-treated samples that exhibited higher concentrations compared to the sonicated urease-treated samples, aligning with our findings. Although the exact origin remains unknown, they postulated that some of the ammonia generated during the urea conversion process in sonication may influence the concentrations of various metabolites. When urea undergoes conversion into ammonia and carbon dioxide, the former may react with acidic species, resulting in the formation of ammonium salts. Since GC-MS cannot detect salts, it may explain the reduced metabolite levels. Furthermore, this study noted that a longer sonication duration (1 h) combined with 200 μ L urease solution is optimal, suggesting a potential avenue for future urease pre-treatment studies involving sonication.

Considering the continued contradictory results regarding urease pre-treatment, researchers should carefully assess its effects and evaluate its potential benefits primarily for specific analyses focused on urea-related measurements, while also ensuring the reliability and validity of the results. It is important to evaluate the trade-off between urea reduction and the potential compromise in metabolite recovery to ensure the most accurate and comprehensive analysis

of the metabolome. If urease pre-treatment is deemed necessary, sonication offers numerous advantages in terms of efficiency, homogeneity, accessibility, avoidance of thermal degradation, and versatility, making it a preferred option to heating.

In terms of heating versus sonication when performing urease pre-treatment, results indicate that sonication is the preferred method to catalyse this reaction. Sonication can be more advantageous than heating due to several reasons, including (i) efficient and rapid mixing and agitation of the sample; (ii) achieving a more homogeneous distribution of the urease enzyme throughout the sample; (iii) facilitating the penetration of the enzyme into cellular compartments or structures, which may be more difficult to access via heating alone; and (iv) avoidance of thermal degradation of sensitive compounds present in the sample as it does not involve significant temperature increases (Kim, 2012; Taufik *et al.*, 2023).

3.12.7. Implications of The Findings

The implications of our findings are significant for clinical, nutritional, and biological investigations focused on examining metabolic abnormalities associated with various pathophysiological conditions. Depending on the specific research objectives, our extraction methods can serve as valuable guidelines. Disease characterisation, for example, typically employs non-targeted metabolomics approaches (such as Method 6) to obtain comprehensive metabolic information about individuals (Drabińska *et al.*, 2020). The data analysis can lead to the development of new hypotheses, facilitating the exploration of previously undiscovered disease mechanisms, host–pathogen interactions and adaptations (Luies & Loots, 2016). When disease characterisation indicates abnormalities in OA or urea profiles, researchers may opt for a more semi-targeted approach (such as Methods 2 or 4, as per our results) or targeted approach (Method 7 with sonication), respectively.

While this study primarily adopts a “discovery mode” process with non-targeted methods, aimed at comparative analysis between different techniques to identify the superior method, it is important to acknowledge the limitations of a non-targeted approach. A targeted metabolomics approach, while providing more accurate quantitative results, focuses only on a predetermined list of metabolites (Gertsman & Barshop, 2018). The optimised extraction method presented here holds promise for application in clinical laboratories, especially where time and resources are limited. This efficient method minimises extraction solvents, making it suitable for high-throughput analysis. Its potential to enhance treatment strategies, overcome diagnostic limitations, and enable prompt medical intervention in disease cases makes it a valuable tool for researchers seeking to optimise urinalysis and make informed decisions aligned with their research objectives.

3.13. CONCLUSION

Since non-targeted GC-MS metabolomics aims to investigate multiple metabolite classes at various concentrations, optimal compound extraction is essential. Five urine sample preparation methods were evaluated in terms of repeatability, metabolome coverage, and metabolite recovery, with the aim of determining which is better for non-targeted GC-MS metabolomics applications. A low-volume DA method outperformed all other methods, consistently ranking first in every validation parameter. This method identified 91 unique peaks, from multiple compound classes, and showed high metabolite recovery. An extraction method that can simultaneously identify all the metabolites in a urine sample is not yet available, mostly due to the complexities and the variety of physiochemical properties of metabolites. Nonetheless, this DA method requires little sample volume, have few analytical steps (which would be more time efficient, cost-effective, and repeatable), and can extract as much of the metabolome as possible, making it ideal for non-targeted urine GC-MS analysis.

Furthermore, the DA approach was optimised by incorporating an additional drying step between two-step derivatization but did not benefit from urease sample pre-treatment. Indeed, we recommend that urease pretreatment should be reserved for specific analyses targeting urea-related measurements. If urease pre-treatment is necessary, sonication is recommended due to its numerous advantages, including efficiency, homogeneity, accessibility, avoidance of thermal degradation, and versatility compared to heating.

3.14. REFERENCES

- Blau, N., Duran, M., Gibson, K.M. & Dionisi-Vici, C. 2014. *Physician's guide to the diagnosis, treatment, and follow-up of inherited metabolic diseases*. 213. Springer.
- Chan, E.C.Y., Pasikanti, K.K. & Nicholson, J.K. 2011. Global urinary metabolic profiling procedures using gas chromatography–mass spectrometry. *Nature protocols*, 6(10):1483-1499.
- Christou, C., Gika, H.G., Raikos, N. & Theodoridis, G. 2014. GC-MS analysis of organic acids in human urine in clinical settings: a study of derivatization and other analytical parameters. *Journal of Chromatography B*, 964:195-201.
- Collino, S., Martin, F.P.J. & Rezzi, S. 2013. Clinical metabolomics paves the way towards future healthcare strategies. *British journal of clinical pharmacology*, 75(3):619-629.
- De Villiers, L. & Du Loots, T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current Metabolomics*, 1(4):306-317.
- Drabińska, N., Młynarz, P., de Lacy Costello, B., Jones, P., Mielko, K., Mielnik, J., ... Ratcliffe, N.M. 2020. An optimization of liquid–liquid extraction of urinary volatile and semi-volatile compounds and its application for gas chromatography-mass spectrometry and proton nuclear magnetic resonance spectroscopy. *Molecules*, 25(16):3651.
- Dunn, W.B. 2008. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Physical biology*, 5(1):011001.
- Esterhuizen, K., Lindeque, J.Z., Mason, S., Van der Westhuizen, F.H., Suomalainen, A., Hakonen, A.H., ... Janssen, M.C. 2019. A urinary biosignature for mitochondrial myopathy, encephalopathy, lactic acidosis and stroke like episodes (MELAS). *Mitochondrion*, 45:38-45.

- Fiehn, O. 2016. Metabolomics by gas chromatography–mass spectrometry: Combined targeted and untargeted profiling. *Current protocols in molecular biology*, 114(1):30.34. 31-30.34. 32.
- Fiehn, O., Kopka, J., Trethewey, R.N. & Willmitzer, L. 2000. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Analytical chemistry*, 72(15):3573-3580.
- Gertsman, I. & Barshop, B.A. 2018. Promises and pitfalls of untargeted metabolomics. *Journal of inherited metabolic disease*, 41:355-366.
- Glicksberg, L. & Kerrigan, S. 2020. Derivatization. *Principles of Forensic Toxicology*:163-175.
- Gupta-Wright, A., Peters, J.A., Flach, C. & Lawn, S.D. 2016. Detection of lipoarabinomannan (LAM) in urine is an independent predictor of mortality risk in patients receiving treatment for HIV-associated tuberculosis in sub-Saharan Africa: a systematic review and meta-analysis. *BMC medicine*, 14:1-11.
- Isa, F., Collins, S., Lee, M.H., Decome, D., Dorvil, N., Joseph, P., ... Fischer, S. 2018. Mass spectrometric identification of urinary biomarkers of pulmonary tuberculosis. *EBioMedicine*, 31:157-165.
- Kanani, H., Chrysanthopoulos, P.K. & Klapa, M.I. 2008. Standardizing gc–ms metabolomics. *Journal of Chromatography B*, 871(2):191-201.
- Khodadadi, M. & Pourfarzam, M. 2020. A review of strategies for untargeted urinary metabolomic analysis using gas chromatography–mass spectrometry. *Metabolomics*, 16:1-14.
- Kim, J., Ahn, J.K., Cheong, Y.E., Lee, S.-J., Cha, H.-S. & Kim, K.H. 2020. Systematic re-evaluation of the long-used standard protocol of urease-dependent metabolome sample preparation. *Plos one*, 15(3):e0230072.
- Kim, J., Choi, K. & Chung, D. 2012. *Comprehensive Sampling and Sample Preparation* (Academic Press.
- Kind, T., Tolstikov, V., Fiehn, O. & Weiss, R.H. 2007. A comprehensive urinary metabolomic approach for identifying kidney cancer. *Analytical biochemistry*, 363(2):185-195.
- Kruve, A., Leito, I. & Herodes, K. 2009. Combating matrix effects in LC/ESI/MS: The extrapolative dilution approach. *Analytica Chimica Acta*, 651(1):75-80.
- Liebeke, M. & Puskás, E. 2019. Drying enhances signal intensities for global GC–MS metabolomics. *Metabolites*, 9(4):68.
- Luies, L. & Loots, D.T. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.
- Madla, S., Miura, D. & Wariishi, H. 2012. Optimization of extraction method for GC-MS based metabolomics for filamentous fungi. *Journal of Microbial and Biochemical Technology*, 4(1):5-9.
- Majors, E. 2013. Sample preparation fundamentals for chromatography. Agilent Technologies. Inc., Nov, 13,
- Mamyrin, B. 2001. Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry*, 206(3):251-266.
- Mobley, H.L. 2001. Urease. *Helicobacter pylori: physiology and genetics*:177-191.
- Mouskeftara, T., Virgiliou, C., Theodoridis, G. & Gika, H. 2021. Analysis of urinary organic acids by gas chromatography tandem mass spectrometry method for metabolic profiling applications. *Journal of Chromatography A*, 1658:462590.
- Nicholson, J.K., Connelly, J., Lindon, J.C. & Holmes, E. 2002. Metabonomics: a platform for studying drug toxicity and gene function. *Nature reviews Drug discovery*, 1(2):153-161.
- Palmas, F., Mussap, M. & Fattuoni, C. 2018. Urine metabolome analysis by gas chromatography–mass spectrometry (GC–MS): Standardization and optimization of protocols for urea removal and short-term sample storage. *Clinica Chimica Acta*, 485:236-242.
- Pasikanti, K.K., Ho, P. & Chan, E. 2008. Gas chromatography/mass spectrometry in metabolic profiling of biological fluids. *Journal of Chromatography B*, 871(2):202-211.
- Patti, G.J., Yanes, O. & Siuzdak, G. 2012. Metabolomics: the apogee of the omics trilogy. *Nature reviews Molecular cell biology*, 13(4):263-269.

- Phiri, M. 2017. *Re-assessment and optimisation of an organic acid extraction method for automation*.
- Qin, Y. & Cabral, J.M. 2002. Review properties and applications of urease. *Biocatalysis and biotransformation*, 20(1):1-14.
- Rokhina, E.V., Lens, P. & Virkutyte, J. 2009. Low-frequency ultrasound in biotechnology: state of the art. *Trends in biotechnology*, 27(5):298-306.
- Schoeman, J.C., Du Preez, I. & Loots, D.T. 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(2):301-311.
- Shulaev, V. 2006. Metabolomics technology and bioinformatics. *Briefings in bioinformatics*, 7(2):128-139.
- Soares da Silva Burato, J., Vargas Medina, D.A., de Toffoli, A.L., Vasconcelos Soares Maciel, E. & Mauro Lanças, F. 2020. Recent advances and trends in miniaturized sample preparation techniques. *Journal of Separation Science*, 43(1):202-225.
- Su, Y., Xia, S., Wang, R. & Xiao, L. 2017. Phytohormonal quantification based on biological principles. *Hormone metabolism and signaling in plants*, 13:431-470.
- Taufik, M., Ardilla, D., Cahyady, B., Alfian, Z., Razali, M., Susilawati, E. & Afniwati, A. 2023. Sonication technique for nicotine extraction from saliva and urine. In. AIP Conference Proceedings. AIP Publishing.
- Webb-Robertson, B.-J., Kim, Y.-M., Zink, E.M., Hallaian, K.A., Zhang, Q., Madupu, R., ... Metz, T.O. 2014. A statistical analysis of the effects of urease pre-treatment on the measurement of the urinary metabolome by gas chromatography–mass spectrometry. *Metabolomics*, 10:897-908.
- Wells, W., Chin, T. & Weber, B. 1964. Quantitative analysis of serum and urine sugars by gas chromatography. *Clinica Chimica Acta*, 10(4):352-359.

CHAPTER 4: SECTION B – METHOD APPLICATION FOR CHARACTERISING THE URINARY METABOLIC EFFECTS OF HIV/TB CO-INFECTION

Parts of this chapter have been published (Appendix C), while other parts are drafted for journal submission:

- Olivier, C. and Luies, L. (2023). WHO Goals and Beyond: Managing HIV/TB co-infection in South Africa. *SN Comprehensive Clinical Medicine*, 5(1), p.251. (<https://doi.org/10.1007/s42399-023-01568-z>)
- Olivier, C. and Luies, L. (2023). Metabolic Insights into HIV, Tuberculosis, and Co-infection: An Untargeted Urinary Metabolomics Approach". *Drafted for intended submission*.

4.1. LITERATURE REVIEW

4.1.1. HIV Virology

Human immunodeficiency virus (HIV) causes acquired immunodeficiency syndrome (AIDS) and is categorised into two strains, HIV-1 and HIV-2 (Nyamweya *et al.*, 2013; Sierra *et al.*, 2005). As lentiviruses, a subgroup of retroviruses (Levy, 1993), both strains share replication mechanisms, transmission modes, and clinical applications — increasing host susceptibility to opportunistic infections. The primary clinical distinction is that HIV-2 progresses to immunodeficiency more slowly than HIV-1, which is also more infectious and transmissible (Sharp & Hahn, 2011). In this study, all references to HIV refer to the HIV-1 strain.

Mature HIV virions, approximately 120 nm in diameter, exhibit a spherical morphology with a lipid bilayer membrane encapsulating a dense, truncated cone-shaped nucleocapsid — the core. This core houses critical components including viral protease, genomic RNA molecules, integrase, reverse transcriptase, and several regulatory proteins such as viral protein U and negative factor protein, along with cellular factors (Hirsch & Curran, 1990; Montagnier & Clavel, 1994).

4.1.2. HIV Basic Pathophysiology

Transmission of HIV occurs through sexual contact, blood exposure, and mother to infant during birth and breastfeeding (Brookmeyer, 1991; Shaw & Hunter, 2012; Stoneburner *et al.*, 1990). The virus infiltrates host cells, converting its RNA genome to DNA via reverse

transcription, integrating into the host's DNA, and commandeering cellular machinery to produce viral proteins and replicate. This replication cycle results in the assembly of new virions capable of further infection (Gomez & Hope, 2005).

The host's response to HIV infection includes an acute phase, a prolonged asymptomatic period, and eventual rapid viral replication coinciding with a decline in CD4+ T cells, which are critical for immune responses. The depletion of these cells over time weakens the immune defence, leading to opportunistic infections and AIDS progression. Clinical monitoring of HIV progression is thus based on CD4+ T cell counts and plasma viral levels (Hernandez-Vargas & Middleton, 2013; Sierra *et al.*, 2005).

4.1.3. TB Bacteriology

Tuberculosis (TB) is caused by *Mtb*, a rod-shaped, aerobic, none-spore-forming bacterium measuring approximately 0.5 µm by 3 µm. A defining characteristic is its complex cell wall, which provides resistance to antibiotics due to its lipid-rich outer layer composed of mycolic acids and polysaccharides. The cell wall's three segments — the plasma membrane, the core, and the outer layer — contribute to *Mtb*'s structural stability and pathogenicity (Brennan, 2003; Grzegorzewicz *et al.*, 2016; Knechel, 2009). The persistence of TB throughout history, with strains found in ancient skeletons, underscores its evolutionary resilience (Smith, 2003).

4.1.4. TB Basic Pathophysiology

Mtb is an airborne pathogen transmitted through aerosols and typically localises in the lung alveoli upon inhalation, although it can also manifest in other body parts (Olivier & Loots, 2011). Post-inhalation, *Mtb* may be cleared, cause latent infection, lead to active disease, or reactivation from latency (Wani, 2013). Within the host, *Mtb* halts phagosome maturation and eludes immune destruction, thriving within macrophages until they expire (Mahamed *et al.*, 2017; Russell, 2001; Warner, 2015). *Mtb*'s replication leverages host resources, such as glucose and lipids, and can continue its life cycle within dead immune cells (Boshoff & Barry, 2005).

Granulomas formed by the immune system encapsulate the bacteria, potentially keeping *Mtb* in a latent state. Progression to active TB, where the bacteria multiply unchecked, can be triggered by factors that weaken the immune system, such as HIV infection, making co-infection a critical concern in regions where both diseases are prevalent (Horsburgh Jr, 2004; Luies & Du Preez, 2020b)

Active TB manifests through local symptoms (at the site of infection) like excessive coughing, fever, night sweats, weight loss which can lead to anorexia, chest pains, haemoptysis

(coughing up blood) and dyspnoea (LoBue & Catanzaro, 1997), and systematic complications including oxidative stress, hyponatraemia, glucose intolerance, hypocholesterolaemia, haematological manifestations, vitamin D deficiency, and an overall altered host microbiota. While the immune system's response can lead to containment in latent TB, active TB results in rapid bacterial replication and spread, with systemic symptoms often being non-specific and difficult to attribute solely to TB. Long-term pulmonary impairment due to TB, influenced by genetic variation in immune response, adds complexity to diagnosis and treatment strategies (Luies & Du Preez, 2020b).

4.1.5. The Syndemic of HIV and TB in South Africa

The intricate relationship between HIV and TB, particularly in South Africa which grapples with an alarming prevalence of both diseases, presents a multifaceted challenge influenced by historical, social, and biological factors. Therefore, it is important to explore the co-dependent nature of these diseases and the challenges to effective control strategies. The ramifications of South Africa's history significantly shaped these pandemics, with TB thriving in impoverished conditions marked by overcrowding and limited health care access. Co-infection intensifies health challenges, further complicating treatment, and management. Despite commendable efforts, achieving the ambitious "End TB" and "End HIV" goals set by the World Health Organization (WHO) for 2030 remains intricate due to socio-economic disparities, health care constraints, and political dynamics. The recent emergence of the COVID-19 pandemic introduced additional complexities, disrupting diagnostic and treatment services. Nevertheless, South Africa's resilience is evident through destigmatisation campaigns, innovative preventative interventions, and significant contributions to global research. The importance of prevalence studies cannot be overstated; accurate data collection and analysis remain pivotal for informed interventions and tracking progress toward WHO goals. It's important to elucidate the need for a unified and comprehensive national response, supported by international collaboration, to effectively address the enduring HIV and TB burden within South Africa. As the nation navigates intricate historical legacies, social factors, evolving health challenges, a crucial question arises: Will South Africa ultimately meet the WHO's elimination goals or at least reduce the number of HIV-related TB deaths? Recognising potential obstacles, it becomes imperative to strategize responses to enable effective syndemic management and advancement towards WHO's objectives.

4.1.5.1. WHO Goals and Beyond: Managing HIV/TB Co-infection in South Africa

The HIV and TB are two of the major infections that continue to be global health threats. Efforts have been undertaken to combat these diseases, with global reports highlighting the urgency of the issue. Global reports indicated there were 1.5 million new HIV infections and 680 000

deaths from AIDS-related causes in 2021 alone. However, substantial challenges persist. An estimated 27.5 million people are currently on antiretroviral therapy (ART) globally. Of the 37.7 million people living with HIV, an estimated 10.2 million is not currently on HIV treatment (HIV/AIDS, 2021). Until 2019, TB was the leading cause of death from a single infectious agent, ranking above HIV/AIDS. While significant progress has been made, the number of patients newly diagnosed with TB between 2020 and mid-2021 in most of the high TB burden countries was significantly below average compared to 2019. An immediate consequence of this can be observed in the number TB deaths in 2020, estimated at 1.3 million people with an additional 214 000 deaths among HIV-positive people (World Health Organization, 2021).

These two infections exert massive pressure on health care systems worldwide, and even more so when combined due to an increase of resources needed to treat co-infected patients. The co-existence of these diseases presents unique difficulties, as HIV and TB have an interconnected relationship which causes severe adverse effects (Bruchfeld *et al.*, 2015; Liebenberg *et al.*, 2021). Notably, HIV/TB co-infection prevalence is highest in the African region with southern Africa exceeding 50% of these cases. This indicates a substantial gap in testing and treatment for co-infected cases, as the amount of people diagnosed with TB that know their HIV status has decreased to 4.2 million (World Health Organization, 2021).

To alleviate the disease burden, the World Health Organization (WHO) compiled a list of high-burden countries regarding global TB cases. While South Africa ranks 8th globally in terms of TB cases, accounting for 3.3% of the total global TB cases, it ranks 1st in HIV cases, with an estimated 7 800 000 people living with HIV — this accounts for 13% of the global HIV cases. In 2020, Africa accounted for a major 25% of the total global TB cases, with South Africa making up a substantial 13.2% of these. Africa was also reported as the region with the highest cases of HIV/TB co-infection (individuals infected with HIV and active TB), South Africa contributing to an astounding 50% of these (HIV/AIDS, 2021; World Health Organization, 2021). Looking ahead to 2030, the WHO continues its resolute stance against TB through the “End TB Strategy”, encompass a holistic strategy that includes (1) integrated patient-centred care to ensure equitable access for individuals affected by TB; (2) implementing inclusive policies and robust systems to address social determinants, strengthen health care systems, and provide financial protection; (3) driving innovation through intensified research; (4) mobilizing communities for awareness and reducing stigma and discrimination; (5) fostering international collaborations for migrant populations; and (6) enhancing data-driven decision-making and adaptive strategies (World Health Organization, 2021). These interconnected elements collectively strive to eliminate the global burden of TB and promote a healthier world.

South Africa plays a pivotal role in advancing global efforts to combat TB by driving research, innovation in treatment, and facilitating clinical trials. This commitment extends to groundbreaking work on HIV preventative treatments, microbicide changes and behavioural changes (Karim & Baxter, 2010; Weinrib *et al.*, 2020). The country has conducted several large-scale clinical trials and contributed research and resources to advance the development of an HIV vaccine (Harris *et al.*, 2022). Another standout achievement is the development of the Xpert *MTB*/RIF test, capable of detecting *Mtb* and Rifampicin resistance (Lawn & Nicol, 2011). Moreover, South Africa's substantial genomic research on TB has unveiled genetic variations linked to increased TB susceptibility, providing important insights into its pathogenesis (Ioerger *et al.*, 2009). This dedication extends to TB vaccine research, exemplified by the South African Tuberculosis Vaccine Initiative (SATVI), which contributes to clinical trials worldwide (Hussey, 2007).

The recent COVID-19 pandemic has introduced significant setbacks to the landscape of addressing HIV and TB in South Africa. The pandemic's impact is multifaceted, leading to increased TB transmission within households due to isolation measures and instilling fear of seeking treatment outside homes, compounded by stigmatisation. This disruption is further emphasized by the negative stigma associated with COVID-19 contraction (Visca *et al.*, 2021). Notably, the pandemic has disrupted essential TB diagnostic and treatment services, significantly impeding overall progress as the COVID-19 pandemic took precedence (World Health Organization, 2021). Consequently, the strides taken in reducing TB mortality rates globally have suffered a setback due to the pandemic's influence (Visca *et al.*, 2021).

4.1.5.2. The Co-Dependent Relationship Between HIV and TB

Africa has the highest rate of HIV/TB co-infection in the world, with a large contribution stemming from South Africa. When two diseases combine, they form an even deadlier syndemic, referring to exacerbated health consequences of the biological interactions observed in multiple individuals within a population when both are present (Singer, 2006). The deadly consequences related to simultaneous infection of a host, has been referred to as "the cursed duet".

In newly HIV diagnosed patients, TB is the most common presenting illness and the most prevalent opportunistic disease among people living with HIV (Tiberi *et al.*, 2017). It was reported that among the total worldwide incident cases of TB, 8% of these was people living with HIV. In Africa, the proportion of TB cases co-infected with HIV was the highest as reported in 2020. Not only are HIV patients more susceptible to contracting TB, but it also lowers treatment success rates. The success rate among co-infected patients is 77% compared to 86% (for first-line TB treatment regimens) in those with only TB (World Health Organization,

2021). Contributing to the low treatment outcomes, is the rapid development of drug resistant strains which complicates treatment (Tiberi *et al.*, 2017). The comparison between HIV and TB deaths in South Africa can be seen in Figure 4.1. This figure also depicts the relationship between the amount of HIV and TB deaths divided by their subsequent incidences expressed as a percentage. This percentage illustrates the correlation between the mortality rate and the incidence of each disease, revealing the number of individuals who succumb to the infection. This percentage has remained relatively steady and experienced a significantly decrease since 2020. An important contributing factor to note is that proper treatment for HIV can substantially extend individuals' lives, which can contribute to this figure (World Health Organization, 2021).

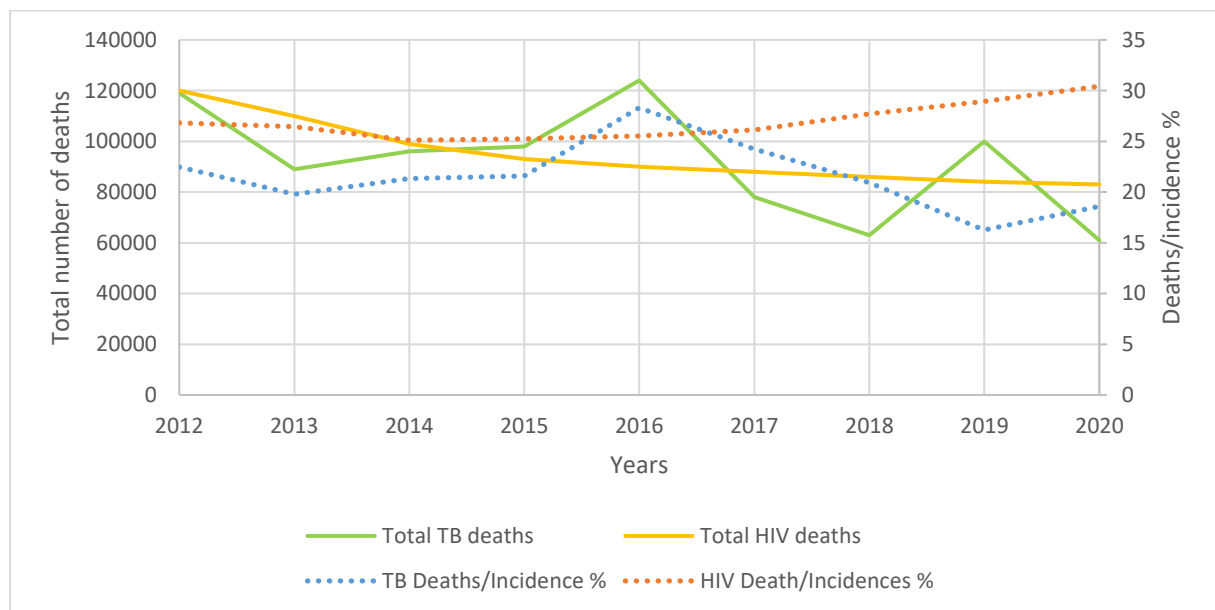


Figure 4.1: The relationship between the total deaths to deaths divided by incidence percentage, for HIV and TB in South Africa. Total TB deaths are indicated with a green line, TB deaths versus incidence percentage is indicated with a dotted blue line, total HIV deaths are indicated with a yellow line, and the HIV deaths versus incidence percentage is indicated with a dotted orange line.

Deaths due to both HIV and TB have been decreasing from around 120 000 in 2012 to around 83 000 for HIV (orange line in Figure 4.1) and 61 000 for TB (green line in Figure 4.1) in 2020. It is, therefore, clear that the prevention of TB deaths has been more successful than the prevention of HIV deaths over the last eight years in South Africa. This is illustrated by the death divided by incidence percentage dotted lines in the graph. TB incidence has decreased to around 18% in 2020, whereas HIV incidence has been rising steadily to around 30% in 2020). Thus, in the case of HIV infections, preventative treatment is relatively successful whereas treatment of the AIDS patients is not that successful. This can also be seen in the number of HIV infections per 1000 people that has been reducing from 8.4% in 2012 to 4.6% in 2020.

The amount of total HIV and TB deaths in South Africa has decreased from 2020 to 2021. Taking the COVID-19 pandemic into account, it can be argued that since HIV-positive individuals' immune systems are more vulnerable, these individuals have a higher chance of contracting COVID-19 and subsequently their death was instead recorded as such and not exclusively an HIV death. The data and records of the causes of death might thus be affected by a global health threat such as the COVID-19 pandemic.

But what causes HIV and TB to have such an intricate relationship and how do they potentiate each other? Although we do not understand the detailed mechanisms of this interaction (Liebenberg *et al.*, 2021), HIV and TB enhance each other causing increased and accelerated deterioration of immunological functions, which when treated poorly, causes an early death for the patient. HIV has been identified in high-burden countries, such as South Africa, as the most important risk factor to develop active TB and in patients with latent TB, it increases the susceptibility to primary infection or the risk of TB reactivation (Kwan & Ernst, 2011). As a patient's CD4+ T cell count decreases, their probability to progress to active TB increases (Tiberi *et al.*, 2017). This increased susceptibility of HIV-infected patients to TB can greatly be attributed to two main mechanisms, namely (1) the increased reactivation of latent TB or (2) the increased risk of contracting *Mtb*. The possibility of both situations being present suggests that these disease states should be closely monitored (Bruchfeld *et al.*, 2015). There are several features of patients with HIV that make them more susceptible to TB infection. One of the main features of HIV/AIDS is the immunosuppression that causes a loss of CD4+ T cells. However, the underlying mechanism of this increased susceptibility is clearly more complicated as it has been shown that TB susceptibility increases soon after HIV infection, which means the CD4+ T cell count is not the only culprit (Badri *et al.*, 2001; Bruchfeld *et al.*, 2015). Another feature of HIV/AIDS, is the sustained immune response induced by the host since a prolonged immune activation leads to an enhanced disease progression, therefore, making these patients more susceptible to TB (Badri *et al.*, 2001). These mechanisms are complex and poorly understood, necessitating further research to enhance comprehension for improved diagnostics and treatment approaches in the HIV/TB co-infected disease state, including the extent to which disease attenuation occurs and the associated metabolic alterations (Liebenberg *et al.*, 2021). However, conventional methodologies employing labour-intensive biochemical assays (Liebenberg *et al.*, 2021) have demonstrated alterations in net protein balance (Paton *et al.*, 2003), hypoalbuminemia (Van Lettow *et al.*, 2003), and similar body composition shifts (Paton & Ng, 2006) in co-infected individuals relative to individuals with HIV infection or TB alone. In 2023, a pioneering study explored the serum metabolic profiles of healthy controls, untreated TB-positive patients, untreated HIV/TB co-infected patients, and HIV/TB co-infected patients on ART. This investigation unveiled substantial shifts in lipid and protein metabolism, alongside translocation of microbial products from the gut to

the bloodstream. The findings imply a synergistic amplification of TB by HIV, contributing to increased inflammation, oxidative stress, ART-induced mitochondrial damage, and detrimental effects on gut health, thereby influencing energy availability. Although ART partially counteracts these trends in co-infected patients, further research is warranted to fully comprehend these changes and their implications (Herbert *et al.*, 2023).

From 2012 to 2020, total TB mortality in South Africa has been showing an overall declining trend (grey line in Figure 4.2). From 2016 to 2020 the TB death percentage based on the total TB deaths reported has been declining for HIV-positive individuals with TB (orange line) compared to HIV-negative individuals (blue line), which could indicate that HIV has been better managed in South Africa since 2016.

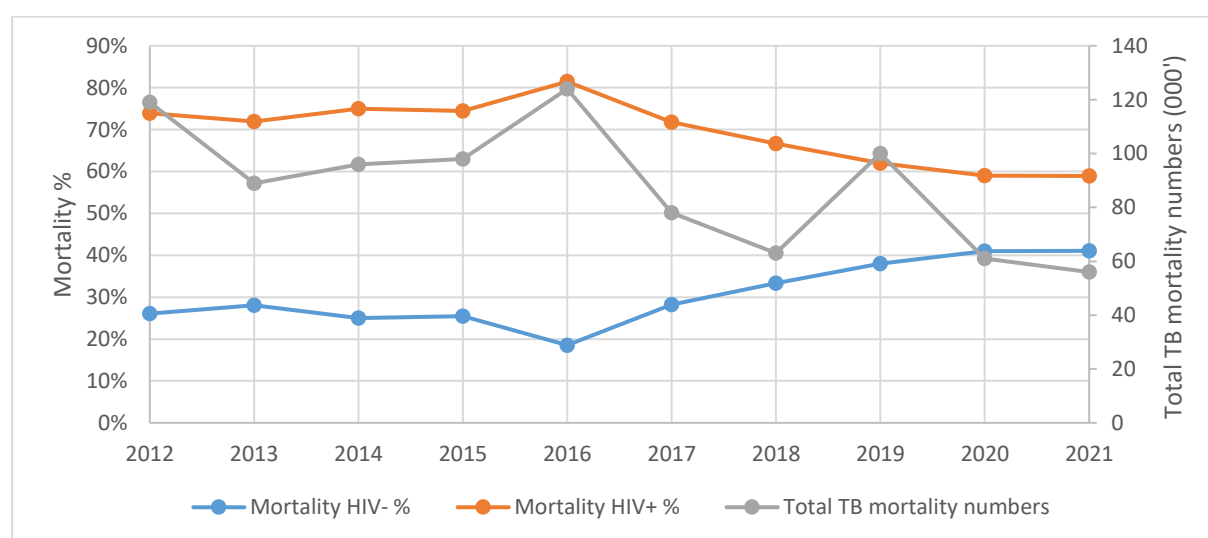


Figure 4.2: Trends in HIV and TB mortality and incidence in South Africa (2012-2021). TB mortality as a percentage of total TB mortality for HIV-positive (orange line) and HIV-negative (blue line) subjects, with total TB mortality numbers (grey line), shown from 2012 to 2021.

4.1.5.3. Social and Biological Dynamics in South Africa That Contribute to HIV/TB Co-Infection

The intertwined histories of both HIV and TB in South Africa have significantly influenced the syndemic challenges of the country today. However, specific social and biological dynamics play pivotal roles in perpetuating the spread of these diseases despite advancements in medical technology. TB has often been named as a “social disease” or “disease of poverty”, alluring to the living conditions and communal behaviours that contribute to its propagation (Lönnroth *et al.*, 2009). Given the substantial proportion of South Africans living in poverty (Mdluli & Dunga, 2022), conditions conducive to HIV and TB transmission prevail. Several studies underscore overcrowding as a notable driver of increased TB transmission (Baker *et al.*, 2008; Pelissari & Diaz-Quijano, 2017). Close-knit living in informal settlements escalates

the risk of both HIV and TB dissemination (Lienhardt, 2001). Furthermore, prevailing social challenges include inadequate food security, limited health care access, substandard living conditions, and a lack of disease awareness (Badri *et al.*, 2001; Lönnroth *et al.*, 2009). The absence of proper education and treatment engagement in these communities exacerbates disease propagation. The intertwining of impoverished circumstances cultivates an environment that promotes HIV and TB transmission (Badri *et al.*, 2001). Moreover, the dual burden of social stigmatisation associated with HIV and TB, which can manifest as fear of discrimination or isolation, as well as misconceptions about the diseases, impedes individuals in communities from seeking essential treatment (Mawar *et al.*, 2005). Effectively addressing disease management may necessitate an integral focus on rectifying the social and biological contexts within South African communities, a vital step towards significant progress.

4.1.5.4. Current Progress Supporting South Africa's Efforts Against HIV and TB

South Africa has recognised the significance of social dynamics in the battle against HIV and TB. Numerous campaigns have been undertaken to confront the stigma and hurdles encountered by communities when seeking treatment. An example of such an effort was initiated by the South African National AIDS Council (SANAC), aiming to counteract persistent social perceptions that continues to fuel stigma and discrimination against people living with HIV. The campaign featured the slogan “I can’t change my HIV status but I can change my attitude” (Peltzer *et al.*, 2012). After this, multiple other campaigns have been introduced in South Africa to educate and destigmatise these diseases within communities.

South Africa has played a significant role in global HIV research. Notably, it has contributed to HIV preventative treatments, including pre-exposure prophylaxis (PrEP). A recent study in South Africa examined the initiation and adherence to oral PrEP in a group of young woman, providing crucial insights on expanding HIV prevention options for this vulnerable group in a high-burden HIV country like South Africa (Mansoor *et al.*, 2022). Further research has also focused on microbicide and behavioural changes (Karim & Baxter, 2010; Weinrib *et al.*, 2020). South Africa has been a site for several large-scale clinical trials of HIV prevention methods with a total of 21 clinical trials initiated up until 2020 (Laher *et al.*, 2020), and even more being initiated since then. Notable successful trials include the HPTN 052 trial, which demonstrated the remarkable efficacy of ART in safeguarding against sexual transmission of the HIV virus. This ground-breaking study revealed that initiating ART in HIV-infected individuals before severe immune system compromise led to a substantial 93% reduction in HIV transmission. These findings underscore the enduring and dependable protective benefits of ART in preventing the spread of the virus through sexual contact (Cohen *et al.*, 2016). Another impactful clinical trial was the ASPIRE trial, a phase III randomized, double-blind, placebo-controlled study evaluating the efficacy of the dapivirine vaginal ring. This study was conducted

across 15 sub-Saharan African sites, including South Africa (Garcia *et al.*, 2021). South Africa also contributed extensively to HIV treatment research over the years, exploring the impact and adherence of South African cohorts to ART administration (Moosa *et al.*, 2019). South Africa has taken crucial steps to identify important barriers relating to treatment regimes, proposing several strategies aimed at enhancing treatment administration in the country to improve outcomes (Coetzee *et al.*, 2004; Maskew *et al.*, 2019; Miller *et al.*, 2010). Efforts towards HIV vaccine development are being pursued through the South African AIDS Vaccine Initiative (SAAVI), established in 1999 in collaboration with government, ESKOM and the Medical Research Council (Tucker & Mazithulela, 2004). South Africa has a high prevalence of both HIV and TB, and co-infection research, considering their metabolic impact, is being thoroughly researched in the country. A recent review examined the complex metabolic effects of HIV and TB and proposed several future objectives (Liebenberg *et al.*, 2021). While a recent HIV vaccine trial, Mosaico phase III, failed in preventing HIV (Venkatesan, 2023), the fight against HIV in South Africa continues unabated. Promising trials for the M72/AS01E TB vaccine are ongoing in South Africa in 2023 (Harris *et al.*, 2022), and an innovative approach using allogeneic hematopoietic stem cell transplantation has shown success in curing HIV, with the involvement of Honorary Professor, Dr Monique Nijhuis (Jensen *et al.*, 2023).

In combating TB, South Africa has made significant contributions, both scientifically and financially. Notable milestones include the development of the Xpert *MTB*/RIF test, which can detect *Mtb* and Rifampicin resistance (Lawn & Nicol, 2011). South Africa participated in a multi-country study evaluating the diagnostic accuracy of this test, collaborating with the Foundation for Innovative New Diagnostics (FIND) (Boehme *et al.*, 2010; Lawn & Nicol, 2011). The active participation of the country in this study, along with several other studies conducted by South African researchers, has played a vital role in the successful rollout of this test. Notably, Professor Nicol and fellow researchers ascertained the accuracy of the test for diagnosing pulmonary TB in a South African Hospital (Nicol *et al.*, 2011). Numerous clinical trials conducted in South Africa have contributed to TB research, including a landmark clinical trial by the Aurum Institute that demonstrated the efficacy of a shorter, simpler TB regime (Salazar-Austin *et al.*, 2019). Significant genomic research on TB has also been conducted in South Africa, identifying genetic variations that increase susceptibility to TB, providing important insights into the pathogenesis of the disease (Ioerger *et al.*, 2009). The country has shown immense dedication to TB vaccine research through the South African Tuberculosis Vaccine Initiative (SATVI), contributing to clinical trials worldwide (Hussey, 2007). Notably, in December 2022, the WHO implemented a new all-oral 6-month treatment regimen for drug-resistant TB in South Africa, showing promising results (Organization, 2022). The Deputy President of South Africa has launched a National Strategic Plan (NSP) to combat HIV and TB, emphasizing breaking down barriers, ensuring equitable and equal access to HIV and TB

services, building resilient health care systems, and establishing revitalised, inclusive, and accountable institutions (Mlangeni, 2023). This plan also acknowledges the impact of pandemics, such as the COVID-19 pandemic, on the fight against HIV and TB, highlighting the need to address burdens caused by such outbreaks. However, it is important to address underlying social economic challenges before the country can achieve the newly outlined goals of the NSP (DeSanto *et al.*, 2023; Foster *et al.*, 2015; Loveday & Zweigenthal, 2011).

While South Africa is actively striving to combat both HIV and TB, the numerous clinical trials and research initiatives, while crucial in the fight against these diseases, do not adequately address the unique social barriers within South Africa that impede progress. The NSP acknowledges these barriers and delineates a well-structured strategy for achieving its four primary goals. However, despite the commendable efforts, there are areas where the practical implementation of these strategies faces challenges. South Africa's history of economic challenges, corruption, and political turmoil presents significant obstacles to achieving the NSP's goals (Schneider *et al.*, 2006). The South African health sector exhibits a range of corrupt practices, including bribery, theft, and subpar service quality (Mantzaris & Pillay, 2021). International research has indicated that in underdeveloped countries, corrupt practices are experienced by up to 80% of the population within the health sector (Holmberg & Rothstein, 2011). With South Africa's national elections coming up in 2024, the successful implementation of the NSP by the current Deputy President may encounter further obstacles in the coming years.

4.1.5.5. Adding Fuel to The Fire

It is evident that the syndemic of HIV and TB has been a major burden for South Africa for a very long time, which was worsened with the recent COVID-19 pandemic. One needs to consider the specific aspects that have contributed to the COVID-19 pandemic adding fuel to the fire, as well as their influence on future progress towards ending HIV and TB.

The WHO 2021 global report estimated that 1.4 million fewer people received care for TB in 2020, a 21% reduction compared to the previous year. The decrease is proposed to be caused by COVID-19-related disruptions such as lockdown implementation and the subsequent difficulty in delivering medicine and administering treatment with restricted access to clinics, hospitals, or treatment centres. Reduced accesses for TB diagnosis and treatment have been one of the biggest effects of COVID-19, ultimately leading to increased TB deaths. TB case detection and reporting have also been negatively affected as COVID-19 case detection and treatment has taken precedence. Several specific disruptions also impacted TB detection, specifically the overall health system capacity being reduced, reduced ability of people seeking health care services during national lockdowns being put in place, fear of contracting the

COVID-19 virus during clinic visits, or stigmas associated with symptoms that are similar between TB and COVID-19, making people less willing to report their symptoms (World Health Organization, 2021).

The HIV pandemic, specifically in South Africa, was also severely impacted by the COVID-19 pandemic. There was a 50% decrease in HIV testing and ART initiation at the beginning of the COVID-19 lockdown, which only marginally increased over the following months (Dorward *et al.*, 2021). Different aspects of the COVID-19 pandemic have contributed negatively towards the progress of ending HIV. Dorward *et al.* (Dorward *et al.*, 2021) suggested a paucity of personal protective equipment, space for physical distancing in clinics, reduced opening times of clinics and staff being redeployed from HIV testing and care to COVID-19 response activities played a crucial role in deterring HIV care. These additional challenges posed by COVID-19 added stress to a country that also faces severe poverty, and issues with overall access to clinics in rural communities. A major challenge for South Africa currently is to reverse the affects COVID-19 had on ending both HIV and TB. The future WHO and UNAIDS reports might shed more light on the impact that COVID-19 had on these two diseases, however, the preliminary data already hints at major setbacks in both the fight against HIV and TB.

When considering South African's TB deaths relative to the reported incidences, and comparing it to that of HIV, for the period of 2012 to 2021, a predicted future relationship can be determined (Figure S3 in Appendix C). If the current trends are followed, the TB number would be higher than expected in 2022/2023, which might reflect the impact of COVID-19 and the overall increase in death experienced by the country. It might, however, still be too soon to see the effects of COVID-19 on HIV as the WHO predict that 2022 figures will reflect this effect better.

4.1.5.6. Are The “End TB” and “End HIV” Goals Out of Reach?

As mentioned previously, the WHO has set targets for ending TB. One of the main goals are to reduce the absolute number of TB deaths by 90% by the year 2030. The number of TB deaths (including HIV-positive and HIV-negative cases) in South Africa has not been declining linearly over the years (Figure S4 in Appendix C). Following the latest data from 2021, if the TB cases are to decline with 90% by 2030, a strengthened control approach is urgently needed to reach the 2030 goal. Similarly, the number of AIDS-related deaths among the South African population has been declining steadily over the years. If the same goals are to be applied to the number of AIDS-related deaths in South Africa, the goal would be a 90% reduction in the death toll by the year 2030. Considering current trends (Figures S4 and S5 in Appendix C), achieving the 2030 goal is highly unlikely — unless an immense, strengthened effort is made and maintained.

The WHO also provides more detailed goals to encourage the fight against TB. Figure 4.3 illustrates the TB death reduction goals set by the WHO (represented by the grey dots) for 2020 (35% reduction), 2025 (75% reduction), 2030 (90% reduction) and 2035 (95% reduction). Although South Africa seems to be on the right track in reducing TB mortality, the country will have to strengthen its overall effort in combatting TB to achieve the WHO goals.

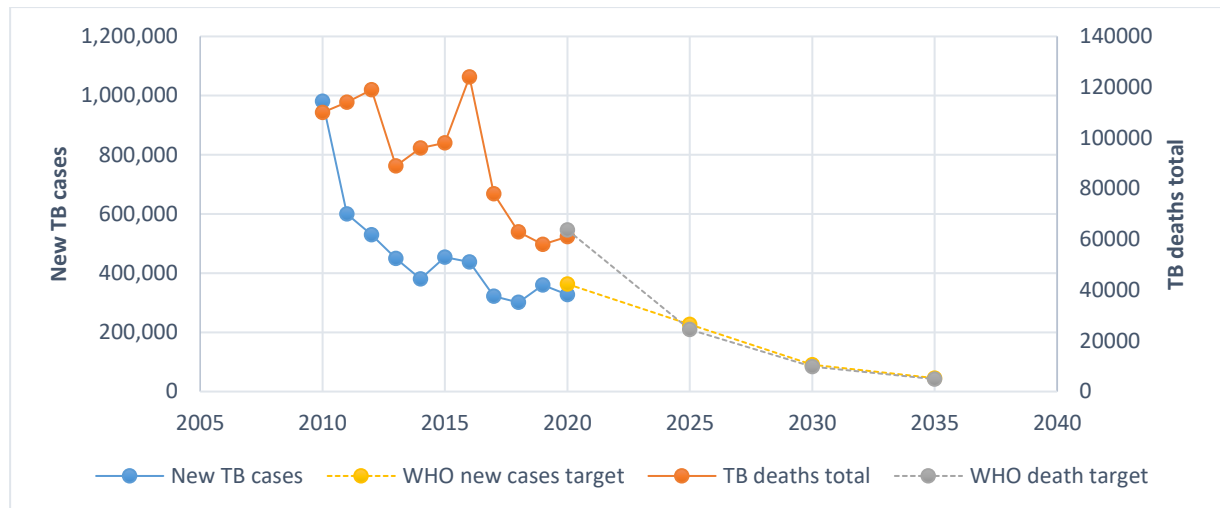


Figure 4.3: Progress towards WHO TB reduction goals in South Africa. New TB cases (blue line) and total deaths (orange line) in South Africa, including the World Health Organization’s global targets (grey line) for 2020, 2025, 2030 and 2035. The new cases targets are indicated with a yellow line.

The goals set by the WHO are meant to motivate countries to end the TB pandemic. As previously mentioned, the WHO’s “End TB Strategy” for 2030 involves a comprehensive approach, including patient-centred care, inclusive policies, innovative research, community mobilization, global collaborations for migrants, and data-driven adaptive strategies to combat TB. Although these goals are usually accompanied by several suggested strategies on how to achieve them (World Health Organization, 2021), they are not personalised to specific counties. Since each country faces its own set of challenges, social dynamics and stigmas should be addressed first before the ultimate goal of ending TB can be achieved.

Although South Africa has made progress in the reduction of TB incidences with the country’s research and NSP implementation, these might not be enough to achieve the “End TB” goals. When considering similar efforts made by the country between 2015 and 2020, the goals were not met. The aim was to reduce TB incidences by 20% and only 9% of the incidences were reduced by 2020. The death rate was also aimed to be reduced by 35% and only 14% reduction was achieved by 2020 (Chakaya *et al.*, 2021). It is thus evident that the long-term “End TB” goals will not be achievable if the problems at hand faced by South Africa is not dealt with first.

The complex interplay of history, society, and health crises has shaped the landscape of HIV and TB in South Africa, impacting the country's battle against these pandemics. Apartheid's shadow still looms, contributing to a disproportionate burden on vulnerable groups and a complex co-infection dynamic. Although progress has been achieved through research and interventions, the persistent high disease rates highlight the urgency of the challenge. The recent COVID-19 pandemic has further complicated matters, undoing previous gains and revealing health care system vulnerabilities. Yet, South Africa's commitment to accurate data via prevalence studies offers hope for informed strategies. Achieving the ambitious "End TB" and "End HIV" goals demands more than data; it requires a coordinated approach, strong health systems, fair access, and robust institutions. While the journey ahead is challenging, past accomplishments driven by resilience demonstrate progress is possible. Confronting historical injustices, overcoming stigmas, and advancing co-infection research require collective effort. Through shared determination, South Africa's pursuit of a healthier, fairer future reflects its resilience and dedication to its people's well-being.

4.2. EXPERIMENTAL DESIGN

While Chapter 2 provides a comprehensive overview of the experimental approach of this study, a more detailed layout specific to Section B is delineated in Figure 4.4.

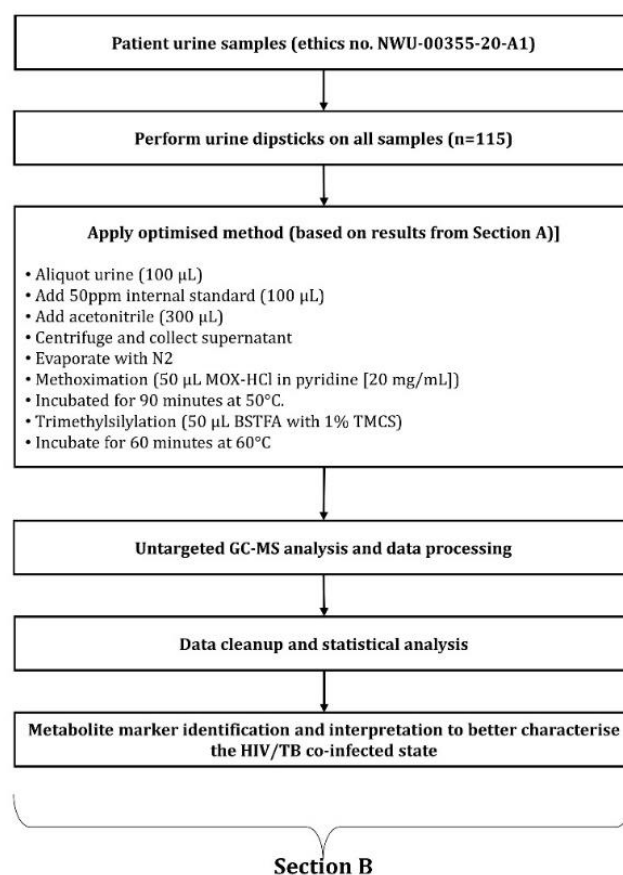


Figure 4.4: Overview of the experimental design for Section B of the study. This figure summarises the workflow from sample collection to data analysis, aiming to achieve the study's objectives.

Within a broader framework, 115 urine samples were collected, however, the groups included in this study were healthy controls (HIV-/TB-/Tn-), TB-positive individuals not receiving treatment (HIV-/TB+/Tn-), HIV-positive individuals not receiving treatment (HIV+/TB-/Tn-), and treatment naïve HIV/TB co-infected individuals (HIV+/TB+/Tn-). A QC sample was constructed by pooling aliquots from each individual sample. The subsequent steps involved direct analysis (as optimised in Section A), derivatization, and GCxGC-TOFMS analysis to get as wide a view of the metabolism as possible, as detailed in Figure 4.4 above.

Machine repeatability was gauged by multiple injections of a single QC aliquot, while analyst repeatability was determined by extracting and injecting multiple QC aliquots. In this sub-study, urine samples collected from healthy controls (n=32), patients with HIV (n=9), patients with TB (n=41), and individuals co-infected with HIV/TB (n=9) were used. Due to the relatively small and non-homogeneous nature of the sample groups, coupled with the lack of comprehensive clinical data — such as absent viral load measurements in many HIV-positive cases — there was insufficient statistical power to pinpoint specific metabolites as definitive biomarkers of HIV/TB co-infection. Consequently, metabolomics was employed in an exploratory capacity to identify broad areas of metabolic alteration or specific classes of metabolites that warrant further detailed investigation in future studies.

The goal was to align the metabolomic profile insights with existing knowledge and to provide a foundation for multidisciplinary approaches required to unravel the complex interactions characterising HIV- and TB infection. Incorporating metabolomics is anticipated to yield valuable insights, necessitating this preliminary work. The results and concepts derived from this study are intended to serve not only metabolomics-focused researchers but also those examining the co-infected state from other scientific perspectives. Thus, these initial findings and their subsequent discussion in the context of existing literature are presented as a cornerstone for future research directions.

4.3. MATERIALS AND METHODS

4.3.1. Sample Collection, Storage, Transport, and Cohort Demographics

The retrospective exploratory study's cohort comprised urine samples from male and female South African citizens, ages of 18 to 69, residing in the Masiphumelele and Ocean View Townships and surrounding areas in Cape Town, Western Cape Province. These samples were collected under standardised protocols (Bi *et al.*, 2020) by the South African Tuberculosis Vaccine Initiative (SATVI) and the Desmond Tutu HIV Centre at the Institute of Infectious Diseases and Molecular Medicine, University of Cape Town (UCT), and cryopreserved at -80°C. Following collection, the samples were transported to the UCT within two hours and

subsequently to North-West University (NWU). Ethical approval was granted for the collection of these samples, which were managed according to UCT's standard operating procedures.

Upon arrival at NWU's Focus Area Human Metabolomics: Laboratory of Infectious and Acquired Diseases, the de-identified samples were stored in bio-freezers equipped with an alarm system, ensuring sample integrity and alerting researchers immediately should any issues arise with the storage units. Access to these samples is restricted to authorized personnel, further safeguarding their quality until the day of analysis.

Geographically confined sampling was employed to minimise the strain variability of *Mtb*, thus making the samples more comparable. Participants were categorised into groups based on serum-confirmed HIV and TB status, diagnosed using GeneXpert for TB. For HIV, dual rapid antibody tests were performed according to South African/WHO guidelines, with positive results leading to appropriate health care referrals. No post-hoc anonymous HIV testing was conducted on the samples.

The study inclusion criteria were stringent, excluding individuals with concurrent diseases, over-the-counter or chronic medication usage, or those pregnant or lactating, to prevent confusing influences on the metabolic profile. In this sub-study, all HIV, TB or co-infected groups were treatment naïve. For these groups, the cohort demographics are elaborated in Table 4.1, which includes age, sex distribution, smoking status, and relevant clinical data such as CD4+ T cell count, and viral load. Notably, clinical data such as T cell count and viral load are only not recorded for HIV-negative individuals and hence this information was not available for all individuals.

Table 4.1: Cohort demographics and clinical data.

	HIV-/TB-	HIV-/TB+	HIV+/TB-	HIV+/TB+
No. of patients (%)	32 (28.1)	41 (36)	9 (7.9)	9 (7.9)
Age (years), mean $\pm$ SD (range)	33.7 $\pm$ 11.2	35.3 $\pm$ 12.2	37 $\pm$ 5.3	33.4 $\pm$ 5.8
Sex, male: female ratio	20:12	10:31	4:5	3:6
Smokers in the group (%)	12 (28.1)	25 (36)	3 (7.9)	3 (7.9)
CD4+ T cell count (cells/mm ³ blood), mean $\pm$ SD	N/A	N/A	Unavailable	N/A
Viral load (copies/mm ³), range*	N/A	N/A	Unavailable	<20–124

Abbreviations: HIV: human immunodeficiency virus; TB: tuberculosis; -: negative; +: positive; SD: standard deviation; CD4: cluster of differentiation 4; NA: not applicable. Groups: HIV-/TB-: healthy controls; HIV-/TB+: untreated patients with active pulmonary TB; HIV+/TB-: untreated patients with HIV; HIV+/TB+: untreated HIV/TB co-infected patients.

**Since the viral load was not available for all patients, the data shown here do not represent the entire group.*

Notably, the sample sizes are relatively small, especially of the HIV-positive (n=9) and HIV/TB co-infected (n=9) cohorts. This is due to the WHO's "test-and-treat" policy, which makes

obtaining untreated HIV and TB samples challenging (World Health Organization, 2021). Thus, this study is rather explorative in nature and aims to guide future research in identifying biomarkers for HIV/TB co-infection rather than discovering them directly.

4.3.2. Reagents and Chemicals

Direct analysis and derivatization methods were performed on the samples using the following reagents: 3-phenylbutyric acid, methoxamine hydrochloride (MOX-HCl), and BSTFA with TMCS. These were acquired from Sigma Aldrich (St. Louis, Missouri, USA). Other chemicals used were acetonitrile, from Burdick and Jackson brand (Honeywell International Inc., Muskegon, USA). Pyridine was also used to dissolve the MOX-HCl and was purchased from Merck (Darmstadt, Germany).

4.3.3. Urine Dipsticks

Before sample extraction was done, Siemens urine dipsticks were used to test the urine samples from each patient group to gather insights into their physiological state and the potential impact on their metabolic profiles. The procedure was as follows: a single test strip was submerged in a well-mixed sample of non-centrifuged urine for exactly 3 s, ensuring full contact of all reagent pads with the urine. After removal, the strip's edges were gently wiped against absorbent paper to eliminate excess urine, taking care not to smear the reagent areas. The test strip was then laid flat, avoiding any overlap of test areas to prevent cross-contamination. After a 60 s wait period, the colours on the reagent pads were compared against the standard colour chart provided on the dipstick container for interpretation.

4.3.4. Creatinine Determination

Creatinine, a breakdown product of creatine phosphate from muscle and protein metabolism, is a normal by-product of urine and serves a critical role in metabolomics studies. It is commonly used as a reference material due to the accepted assumption that its concentration in urine reflects the overall concentration of the sample (Ryan *et al.*, 2011). In the context of a metabolomics study, the concentration of a specific metabolite is divided by the creatinine concentration determined of that urine sample. This ratio is typically reported as $\mu\text{mol}/\text{mmol}$ creatinine, providing a normalised measure of the metabolite concentration relative to creatinine (Bouatra *et al.*, 2013). This normalisation is vital as it accounts for variations in urine concentration among different samples, ensuring that the metabolite levels are compared on a consistent basis. For this study, the creatinine values were accurately determined by the Centre for Human Metabolomics (CHM) at the North-West University (Potchefstroom Campus) using a creatinine enzyme kit (Thermo Fisher Scientific; reference number 981845) and an

Indiko Clinical Analyzer, Type 863 (Thermo Fisher Scientific, Massachusetts, USA). This ensures the reliability of the metabolite concentration measurements, an essential aspect of the study's analytical methodology.

4.3.5. Sample Extraction and Derivatization

Based on the optimised procedures established in Section A, sample extraction was carried out in accordance with the protocol specified in Section 3.9.6 ("Method 6"). Briefly, to each 100 μ L urine sample, a 50ppm internal standard (100 μ L) was added, followed by the addition of ice-cold acetonitrile (300 μ L). The mixture was then centrifuged at room temperature (15 000xg for 5 minutes), and the resulting supernatant was carefully transferred into a GC-MS vial. The samples were dried under a stream of nitrogen gas at 37°C. For oximation, 50 μ L of MOX-HCl in pyridine (20 mg/mL) was added, and the samples were incubated for 90 minutes at 50°C. After a second drying step under the same conditions, silylation was performed using 50 μ L of BSTFA with 1% TMCS for 60 minutes at 60°C. The prepared samples were then placed into glass inserts, transferred to GC vials, and securely capped for analysis (Olivier *et al.*, 2023).

4.3.6. GCxGC-TOFMS Analysis

GCxGC-TOFMS analysis was conducted as indicated in Section 3.10. In summary, following thorough instrument preparation, samples were analysed in seven randomised batches using a Pegasus 4D GCxGC-TOFMS system with an Agilent 7890 GC, employing a dual-column setup for chromatographic separation. Samples were injected using a 1:10 split ratio with purified helium as a carrier gas, set at a constant flow of 1.4 mL/min. The temperature programmes for both primary and secondary ovens ramped from 70°C to 300°C. Cryomodulation and precise temperature control were utilised throughout the run, which lasted approximately 24 minutes per sample. A solvent delay was applied before mass spectral data collection, ranging from 50–800 m/z, with optimised detector settings. The temperature programme is available in Figure 4.5.

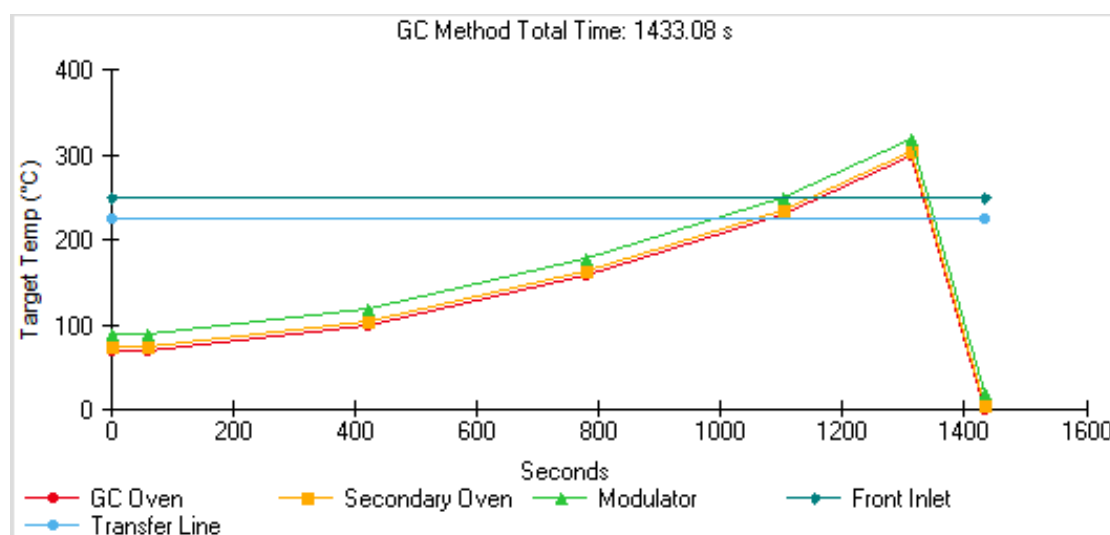


Figure 4.5: GC ramping temperature programme. The temperature profile of the GC run, starting at 70°C and methodically increasing to the final temperature of 300°C, outlining the precise conditions for the separation of compounds.

4.3.7. Data Processing and Data Clean-Up

Data processing and cleanup are essential steps to ensure that high-quality data can be interpreted in metabolomics studies. A summary of the steps applied to the data is outlined in Figure 4.6 below.

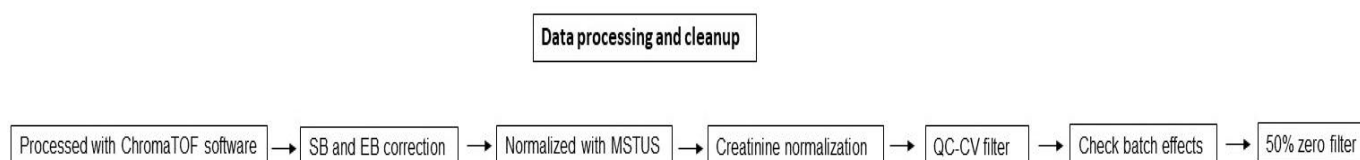


Figure 4.6: Overview of data processing and cleanup steps in metabolomics studies. This figure illustrates the essential steps of data processing and cleanup in metabolomics research. It shows the workflow from System Blank (SB) and Extraction Blank (EB) usage, through mass spectral total useful signal (MSTUS) analysis, to quality control-coefficient of variance (QC-CV) assessment. The role of LECO Corporation's ChromaTOF software in deconvolution, peak alignment, and peak identification is highlighted, showcasing the transition from raw data to analysed metabolites.

After GCxGC-TOFMS analysis, the data was processed using LECO Corporation's ChromaTOF software, version 4.32. This software is a widely accepted mass spectrometer data system for acquiring, processing, and reporting data on analysed metabolites. Its key functionalities include deconvolution, peak alignment, and peak identification. Peak co-elution, a common occurrence during GCxGC-TOFMS analysis where multiple compounds elute at the same retention time, is addressed using deconvolution. This process separates co-eluting peaks based on their distinct spectra, generating a pure spectrum for each component and

enabling more accurate data analysis (Du & Zeisel, 2013). Additionally, the software facilitates peak alignment, aligning corresponding peaks across multiple sample runs, which is crucial for consistent peak identification (Issaq *et al.*, 2009). For peak identification, ChromaTOF utilises a comprehensive spectral library. In this investigation, peaks were identified using both in-house libraries, compiled from previously injected standards, and the NIST (National Institute of Standards and Technology) mass spectral library. The software's strength lies in its robust spectral matching capability, enabling reliable identification of eluted peaks based on spectral patterns, rather than relying solely on mass accuracy. In non-targeted metabolomics approaches, three important procedures are followed: converting spectral data into identifiable metabolite information, analysing these metabolites, and interpreting the significance of identified metabolites in the context of the study (Du & Zeisel, 2013).

Hereafter, data cleanup was performed before statistical analysis, adhering to standard practices in metabolomics research. Initially, system blank and extraction blank corrections were made. Both system blanks and extraction blanks are injected alongside samples to identify contaminants by comparing compounds in the samples to those in the blanks. For each compound, its average concentration in the samples is divided by its average concentration in the system blanks to calculate a ratio. If this ratio is 20 or less, the compound is considered a contaminant and is removed from the dataset. Conversely, if the ratio is greater than 20, indicating that the compound's concentration in the samples is significantly higher than in the system blanks, it is retained in the dataset. Peaks in the extraction blanks are assessed similarly, however, if this ratio is 3 or less, the peak is removed from the dataset (Broadhurst *et al.*, 2018).

Next, the data was normalized using two approaches: mass spectral total useful signal (MSTUS) and creatinine normalization. MSTUS normalization was applied to adjust for variations in sample “weight” or urine concentration, thereby standardising the dataset for comparative analysis (Chetwynd *et al.*, 2016). Additionally, creatinine normalization was employed as a complementary measure to correct for individual variations in urine dilution that MSTUS alone might not fully address. Creatinine, a relatively stable metabolic product in urine, serves as a reliable marker to normalize urine volume and concentration across different samples, ensuring a more accurate representation of metabolite concentrations. The combined use of MSTUS and creatinine normalization offers a comprehensive approach, enhancing the robustness and reliability of our metabolomic data analysis.

Thereafter, a 50% quality control-coefficient of variation (QC-CV) filter (Schoeman *et al.*, 2012) was applied by eliminating compounds with CV values greater than 50% indicating excessively fluctuating compounds. The data was also checked for any batch effects. Batch effects can significantly reduce statistical power and adversely affect reproducibility and repeatability of

results if not properly addressed. There are two types of batch effects that can occur in datasets: between-batch effects and within-batch effects. The former may be caused by systematic variations in the instrumental response, leading to data being segmented into several analytical batches. Within-batch effects occur, on the other hand, can arise from variations due to inlet interface contamination or changes in column performance, also resulting in a split in the data according to the batches (Sánchez-Illana *et al.*, 2018). To check for these effects, principal component analysis (PCA) was performed on the QC samples. This statistical method is effective for detecting batch-related variances, as it visualises the clustering of data points. As shown in Figure 4.7, the PCA scores plot indicated no significant batch effects when considering the QC samples, demonstrating a consistent instrumental response and stable column performance throughout the analysis.

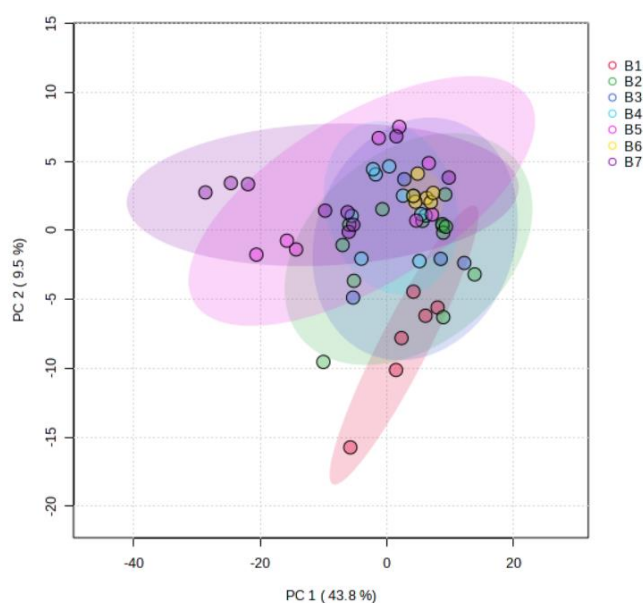


Figure 4.7: Principal component analysis scores plot indicating absence of batch effects. This plot visually confirms that the QC samples do not exhibit any significant between-batch or within-batch effects, demonstrating consistent instrumental response and stable column performance throughout the analysis.

After inspecting for batch effects, the relevant samples for this metabolomics study were isolated from the dataset. A 50% zero filter was then applied to the dataset to remove compounds not present in at least 50% of the total samples. The choice of a 50% threshold is based on a balance between maintaining sufficient data for robust analysis and removing spurious or infrequently detected compounds that could introduce variability or noise into the results (Smuts *et al.*, 2013). This threshold helps ensure that only metabolites consistently present across most samples are included in the analysis, thereby enhancing the reliability of the findings.

4.3.8. Statistical Data Analyses

For the statistical analysis of the processed data in this investigation, MetaboAnalyst (version 5.0), a comprehensive web-based server for metabolomics studies, was used (Pang *et al.*, 2021). This platform offers a range of statistical tests and methods to discern significant differences between various groups. MetaboAnalyst supports both parametric and non-parametric tests. Parametric tests are based on assumptions about the distribution shape and parameters of the underlying population, whereas non-parametric tests make minimal assumptions about the population distribution's shape or parameters (Hoskin, 2012).

The data underwent normalization through log-transformation and autoscaling to approximate a normal distribution, a prerequisite for many statistical tests (Field, 2013). However, it is important to recognise that log-transformation does not invariably result in a perfectly normal distribution, particularly in datasets with small sample sizes or inherent non-normal features. Given our study's relatively small sample size, which can impede the effectiveness of transformation in achieving true normality and increase the likelihood of skewness and outliers, careful consideration was given to the choice of statistical tests. In our analysis, both univariate and multivariate analyses were applied. Univariate analysis, which assesses a single variable at a time, was conducted using the Wilcoxon signed-rank test. Although typically used for paired data, in this study, the Wilcoxon signed-rank test was employed to compare the distribution of metabolic profiles between each of the diseased groups (i.e., HIV-positive, TB-positive, and HIV/TB co-infection) and healthy controls independently (Hollander *et al.*, 2013). This approach was chosen due to its effectiveness in handling non-parametric data and its suitability for datasets that may not conform to a normal distribution, especially in the context of our study's small sample size. Also, the non-parametric Kruskal-Wallis test, an alternative to one-way analysis of variance (ANOVA), was used for comparing multiple independent groups, appropriate for our study's small sample size and potential non-normal or ordinal data distribution (Taylor, 2023).

Multivariate analyses, assessing multiple variables simultaneously (Saccenti *et al.*, 2014), included PCA and partial least squares discriminant analysis (PLS-DA). These methods provide an integrated view of the data, highlighting patterns and relationships not apparent in univariate analysis. PCA was utilised to condense the multivariate data into principal components, simplifying the complex data structure into a more manageable form and aiding in identifying patterns, trends, or outliers (Eriksson *et al.*, 2013). PLS-DA, as a discriminant analysis method, useful for classifying observations into predefined groups based on their metabolic profile (Eriksson *et al.*, 2013; Saccenti *et al.*, 2014).

4.4. RESULTS

4.4.1. Machine and Analyst Repeatability

Machine and analyst repeatability were determined from the QC results obtained during analysis. To assess machine repeatability, a single QC aliquot was extracted and injected multiple times, and the repeatability was quantified using standard metrics such as CV, ensuring alignment with accepted analytical standards. For analyst repeatability, multiple QC aliquots were extracted and each injected once. In Figure 4.8(a), the overlay of a single QC sample injected three times shows identical chromatograms, indicating sufficient machine repeatability. This consistency is crucial for the reliability of the GC-MS analysis. When analysing Figure 4.8(b), the overlay of QC samples from each batch, extracted on seven different days, shows identical chromatograms, which indicates good and acceptable analyst repeatability for this study.

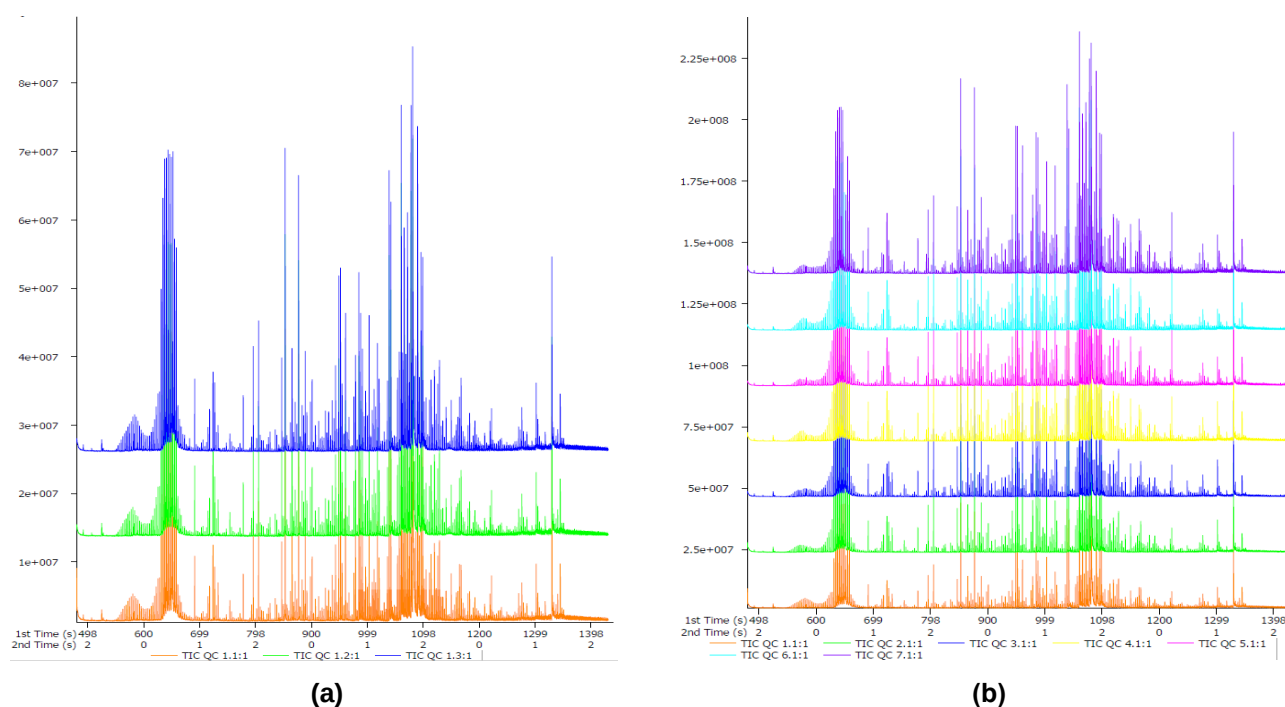


Figure 4.8: Chromatogram overlays demonstrating machine and analyst repeatability. Here, (a) shows the consistency in chromatograms from a single QC sample injected thrice, indicating machine repeatability, while (b) displays chromatograms from QC samples across seven days, highlighting consistent analyst performance. A baseline offset of 1×10^7 and 2.5×10^7 was chosen for visualisation, respectively.

4.4.2. Urine Dipsticks

Urine dipsticks are important for understanding some metabolic changes that could be related to the diseases studied and should be considered in conjunction with other results. The results

of these dipsticks, summarised in Table 4.2, show variations in parameters such as leukocytes, nitrites, urobilinogen, and others across the different patient groups. The layout of the urine dipstick with each test parameter is also laid out in Figure 4.9.

Table 4.2: Summary of urine dipsticks results, comparing the presence and levels of various urinary markers across different sample groups.

	HIV-/TB-	HIV-/TB+	HIV+/TB-	HIV+/TB+
Number of samples	32	41	9	9
Blood presence	Detected (6%)	Detected (12%)	Not detected	Detected (11%)
Bilirubin	Positive (6%)	Positive (7%)	Negative	Positive (11%)
Urobilinogen	Normal	Normal	Normal	Normal
Ketones	Negative	Negative	Negative	Negative
Glucose	Positive (3%)	Positive (7%)	Positive (11%)	Negative
Proteins	Detected (6%)	Detected (21%)	Not detected	Detected (44%)
Nitrites	Positive (3%)	Positive (2%)	Negative	Negative
Leucocytes	Detected (15%)	Detected (7%)	Detected (11%)	Detected (11%)
pH average	6.22	5.91	6.21	5.94
Specific Gravity average	1.02	1.02	1.02	1.02

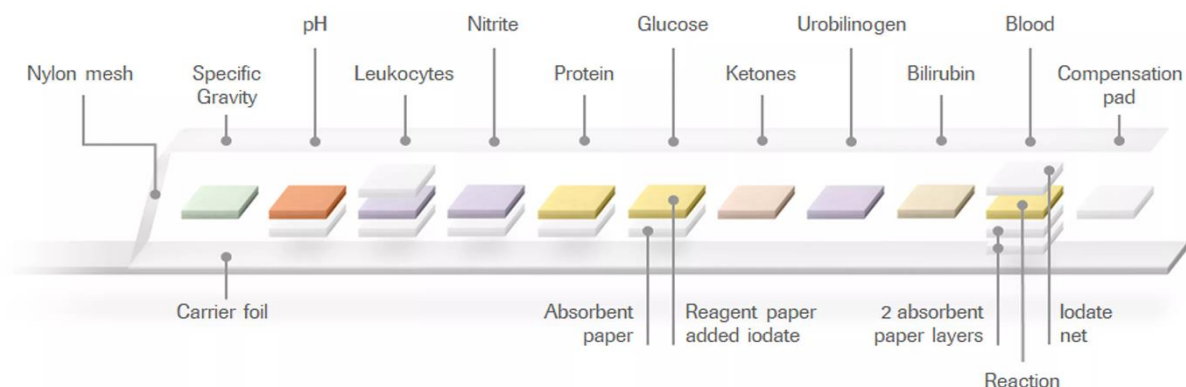


Figure 4.9: Urine dipsticks layout. This image (from Roche) shows the layout of the urine dipsticks used in the study, aiding in understanding the methodology and interpretation of results.

Blood presence was noted in three groups except for the HIV-positive group. Notably, certain bacteria, like *Mtb*, can cause false-positives due to their hydroperoxidase activity. This could explain the blood presence in the TB group, while in the control group, it might be indicative of a urinary tract infection. It is also important to consider that in menstruating individuals, menstrual blood can contaminate urine samples, leading to false-positive results for blood. Therefore, the timing of sample collection in relation to menstrual cycles should be considered when interpreting these results (Boyd & Barratt, 2011).

Urobilinogen levels were normal across all groups, typically a sign of liver health (Roberts, 2007). However, the presence of bilirubin in three of the groups contradicts this, suggesting possible liver or bile duct issues, potentially linked to the primary infections of TB or HIV.

Ketones, indicating fatty acid metabolism when the body cannot use carbohydrates, were not detected in any sample group, suggesting normal carbohydrate metabolism. However, these samples were not collected in a fasting state, which can influence ketone levels. Typically, fasting can lead to increased ketone production as the body starts to utilise fat for energy in the absence of available carbohydrates. To this end, while the presence of glucose in three of the sample groups could indicate conditions like diabetes mellitus or hyperglycaemia, since the individuals were not fasting, this can impact glucose levels in the urine. Non-fasting samples can have elevated glucose levels due to recent carbohydrate intake. Therefore, the presence of glucose in these samples should be interpreted with caution, considering dietary intake prior to the sample collection (Boyd & Barratt, 2011).

Protein detection in urine, done via colour change with tetrabromophenol blue, was noted in three groups except for the HIV-positive group. Protein presence, or proteinuria, could point to various conditions ranging from stress and certain cancers to renal dysfunction (Boyd & Barratt, 2011).

Nitrites, converted from nitrates normally excreted by the kidneys, were found in the control and TB-positive groups. This could indicate the presence of bacteria, potentially signifying urinary tract infections (Boyd & Barratt, 2011; Roberts, 2007). Particularly in the TB-positive group, the tuberculosis bacteria might be contributing to this conversion.

Leukocytes (white blood cells) were detected in all sample groups. Their presence in the urine samples typically indicates infection or inflammation (Roberts, 2007). For patients with HIV or TB, this could be attributed to the diseases themselves. However, in the healthy group, it might suggest unrelated inflammatory diseases, like kidney inflammation or urinary tract infections.

The pH levels of all groups fell within the normal human urine range (5–6.5). However, urine pH can be influenced by dietary habits (e.g., proteins or acidic food would decrease the pH) and medication intake (e.g., anti-epileptic medication would increase the pH), limiting its use as a definitive diagnostic tool (Boyd & Barratt, 2011).

Urine specific gravity was within the normal range (1.003–1.030) (Roberts, 2007) in all groups and therefore did not indicate any particular disease.

It is important to note that the interpretation of these dipstick results can be subjective and may vary based on the analyst's experience and the testing environment. Factors such as lighting conditions and colour perception can influence the reading of these tests. Therefore, these

results should be considered as indicative rather than definitive, and where possible, corroborated with additional diagnostic methods.

4.4.3. Univariate Testing

4.4.3.1. Kruskal-Wallis Test: Statistical Comparison of All Four Groups

The Kruskal-Wallis test, as previously described, was employed to compare differences across the four cohort groups: TB-positive, HIV-positive, co-infected, and a healthy control group. While this test effectively determines whether statistically significant differences exist among these groups, it can identify that at least two groups differ from each other, but it does not indicate which specific groups these are. To discern this, a post-hoc test is required.

In our study, the Kruskal-Wallis test identified 15 metabolites (visualisation shown in Figure 4.10) differing significantly between the groups. These urinary metabolic compounds often result from a complex interplay of various factors, reflecting combined metabolic pathways rather than a single cause.

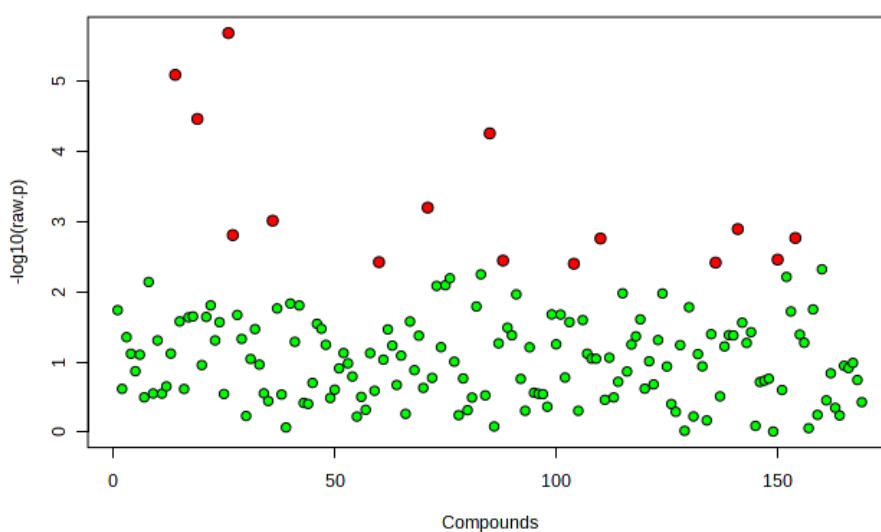


Figure 4.10: Kruskal-Wallis test visualisation. This figure illustrates the outcomes of the Kruskal-Wallis test, where red dots represent statistically significant metabolites, and green dots denote metabolites not found to be significant.

The significant metabolites, detailed in Table 4.3, offer insights into the multifaceted nature of the metabolic alterations associated with each condition and aids the exploratory and hypothesis-generating aim of the current study. It is important to note that in the statistical evaluation of these metabolites, post-hoc analysis results were consistently noted as “not applicable” for all comparisons. This may be attributed to factors such as the small sample size and the specific characteristics of our non-parametric dataset, which can limit the applicability or interpretability of post-hoc tests in certain contexts. Therefore, the data in Table 4.3 focuses

on the group with the highest average concentration for each significant metabolite, as this was the most consistently informative metric derived from our analysis.

Table 4.3: Significant metabolites identified by the Kruskal-Wallis test.

	p-value	FDR (0.05)	Group with highest average concentration
Quinolinic acid	0.001	0.022	Co-infected
Diaveridine	0.002	0.030	Co-infected
Gluconolactone	0.004	0.045	Healthy controls
Hippuric acid	<0.001	<0.001	Healthy controls
Citric acid	0.001	0.028	HIV-positive
Glycine	0.004	0.045	HIV-positive
Hydracrylic acid	0.001	0.030	HIV-positive
Meso-Erythritol	0.002	0.030	HIV-positive
Methylmalonic acid	0.004	0.045	HIV-positive
Pentanedioic acid	0.002	0.030	HIV-positive
Propanoic acid	<0.001	0.001	HIV-positive
2-Deoxyribolactone	0.004	0.045	HIV-positive
2-Methylene-butane-1,4-diol	0.004	0.045	HIV-positive
2,3-Butanediol	<0.001	0.002	HIV-positive
2,3-Dihydroxybutanoic-acid	<0.001	0.002	HIV-positive

Abbreviations: FDR: false discovery rate.

Conclusions drawn from these results are based on prior literature and knowledge of metabolism. The intricate relationships among these metabolites are further explored in the discussion section. For a comprehensive view of the data distribution and variability across all groups, a detailed analysis of these metabolites, including box and whisker plots, is presented in Appendix D (Table A1). These additional visualisations provide a nuanced view of the entire range of metabolite concentrations, facilitating a deeper interpretation that encompasses both high and low concentration values, thus allowing us to draw meaningful conclusions about the metabolic alterations associated with each condition.

4.4.3.2. Wilcoxon Signed-Rank Test: Pairwise Statistical Comparison

While the previously mentioned Kruskal-Wallis test identified significant metabolites indicating differences among the four cohort groups, it did not specify which particular groups differed (Hecke, 2012). To address this limitation and identify specific group differences, pairwise comparisons were conducted using the Wilcoxon signed-rank test. This non-parametric test is particularly suitable when data is not normally distributed (McKight & Najab, 2010).

In this study, five pairwise comparisons using the Wilcoxon signed-rank test were conducted to determine specific differences between groups, with a p-value threshold set using a false discovery rate (FDR) of 0.05 for significance:

1. HIV-positive versus healthy controls (Objective 4b): One significant metabolite, pentanedioic acid, was identified (p-value: 0.0002; FDR: 0.034), indicating an upward trend in the HIV-positive group.
2. TB-positive versus healthy controls (Objective 4a): Twenty-two significant metabolites were detected (listed in Table 4.4), highlighting various metabolic differences.
3. Co-infection versus healthy controls (Objective 4c): No significant metabolites detected.
4. Co-infection versus HIV-positive (Objective 4c): No significant metabolites detected.
5. Co-infection versus TB-positive (Objective 4c): No significant metabolites detected.

Table 4.4: Trend in the TB-positive group when compared to healthy controls.

	P-value	FDR	Trend in TB
2-Hydroxycyclohexane carboxylic acid	0.001	0.023	↑
3-Phenyllactic acid	0.006	0.047	↑
Biphenyl	0.006	0.047	↑
Butanediol	<0.001	0.004	↓
Citric acid	0.002	0.024	↓
Diaveridine	0.002	0.029	↑
2,3-Dihydroxybutanoic acid	0.006	0.047	↓
2,3-Dihydroxybutanoic acid	<0.001	0.002	↓
Ephedrine	0.006	0.048	↑
Gluconolactone	0.001	0.022	↓
Hippuric acid	<0.001	<0.001	↓
Hydracrylic acid	0.001	0.022	↓
Meso-Erythritol	0.001	0.014	↓
Methylene-butane	0.001	0.014	↑
Methylsuccinic acid	0.005	0.047	↓
Octanoic acid	0.003	0.031	↓
Propanoic acid	<0.001	0.001	↓
Quinolinic acid	<0.001	0.004	↑
Ribitol	0.003	0.031	↓
Threonic acid	0.005	0.047	↓
Tyrosine	0.003	0.031	↑
Valeric acid	0.003	0.031	↓

Abbreviations: FDR: false discovery rate.

These additional pairwise comparisons using the Wilcoxon signed-rank test were conducted to complement the limited number of variable importance in projection (VIPs) identified in the Kruskal-Wallis test. This approach was chosen to provide a more comprehensive analysis and enrich our understanding of the data, despite some degree of redundancy, thereby enhancing the robustness of our results.

4.4.3.3. Summary

This exploratory proof-of-concept study takes a holistic approach in analysing metabolite variations, combining results from the Kruskal-Wallis and Wilcoxon signed-rank tests to compile a list of noteworthy metabolites (Table 4.5). This method differs from more targeted studies as it encompasses a broader range of metabolites, designed to guide future research. The list includes metabolites highlighted across all four groups (determined by the Kruskal-Wallis test) and those significant between specific group pairs (identified by the Wilcoxon signed-rank test), detailing their observed trends and the statistical tests underpinning their significance.

Table 4.5: Significant metabolite summary.

Metabolite*	<u>Kruskal-Wallis test results</u> <u>(Table 4.3):</u> Group with highest average concentration	<u>Wilcoxon signed-rank test</u> <u>result (Table 4.4):</u> Trend and comparison
Amino Acids and Related Metabolites		
3-Phenyllactic acid	-	TB ↑ versus HC
Glycine	HIV-positive	-
Hydracrylic acid	HIV-positive	TB ↓ versus HC
Quinolinic acid	Co-infected	TB ↑ versus HC
Tyrosine	-	TB ↑ versus HC
Carbohydrates, TCA cycle, and Energy Metabolism		
Citric acid	HIV-positive	TB ↓ versus HC
Treonic acid	-	TB ↓ versus HC
Meso-Erythritol	HIV-positive	TB ↓ versus HC
Ribitol	-	TB ↓ versus HC
Gluconolactone	Healthy controls	TB ↓ versus HC
Dicarboxylic Acids and Lipids		
2-Methylene-butane-1,4-diol	HIV-positive	-
Methylene-butane	-	TB ↑ versus HC
Methylmalonic acid	HIV-positive	-
Octanoic acid	-	TB ↓ versus HC
Pentanedioic acid	HIV-positive	HIV ↑ versus HC

Propanoic acid	HIV-positive	TB ↓ versus HC
Gut Microbiota Imbalance		
Hippuric acid	Healthy controls	TB ↓ versus HC
Valeric acid	-	TB ↓ versus HC
2,3-Dihydroxybutanoic acid	HIV-positive	-
2,3-Butanediol	HIV-positive	-
DNA Damage and Oxidative Stress		
2-Deoxyribolactone	HIV-positive	-
Dietary Origin		
2-Hydroxycyclohexane-carboxylic acid	-	TB ↑ versus HC
Diaveridine	Co-infected	TB ↑ versus HC
Exogenous Compounds		
Ephedrine	-	TB ↑ versus HC
Other / Unclassified		
Biphenyl	-	TB ↑ versus HC

Abbreviations: HIV: human immunodeficiency virus; TB: tuberculosis; HC: healthy controls; ↑: increased; ↓: decreased.

*Compound names align with Metabolomics Standards Initiative (MSI) Level 2 identification.

In our study, the identification of metabolites using GC-MS primarily aligns with Level 2 of the Metabolomics Standards Initiative (MSI) scoring system. This categorisation indicates that metabolites were putatively annotated based on spectral similarity to known compounds in public and in-house libraries, without direct comparison to chemical reference standards. Consequently, while there is a reasonable degree of confidence in the identification of these compounds, we acknowledge the inherent limitations of this approach and the possibility of misidentification. This uncertainty is an inherent aspect of metabolomics research, especially in analyses that do not involve verification with authentic standards. Therefore, the identified metabolites should be considered with an understanding of these potential limitations.

4.4.4. Multivariate Testing

4.4.4.1. PCA and PLSDA:

In this study, PCA, an unsupervised method, and PLSDA, a supervised method, were utilised. These tests were performed to analyse the intricate relationships between the different groups in the study, namely (1) HIV-positive and (2) TB-positive versus HIV/TB co-infection. The outcomes of these tests are presented in Figure 4.11.

PCA was employed to explore the inherent structure of the data. However, it showed limited separation between the groups as indicated in Figure 4.11(a) and (b), leading to its limited

utility in further metabolite marker selection. In fact, the PCA power, being below 0.5, did not provide a level of significance that warranted further exploration. PLS-DA, in contrast, displayed clearer group separation, particularly useful for classification purposes, as depicted in Figure 4.11(c) and (d). For PLS-DA, metabolites with a VIP score greater than 1 are considered significant. However, this determination is subject to the outcomes of cross-validation and permutation testing, which are essential to assess the reliability of these findings.

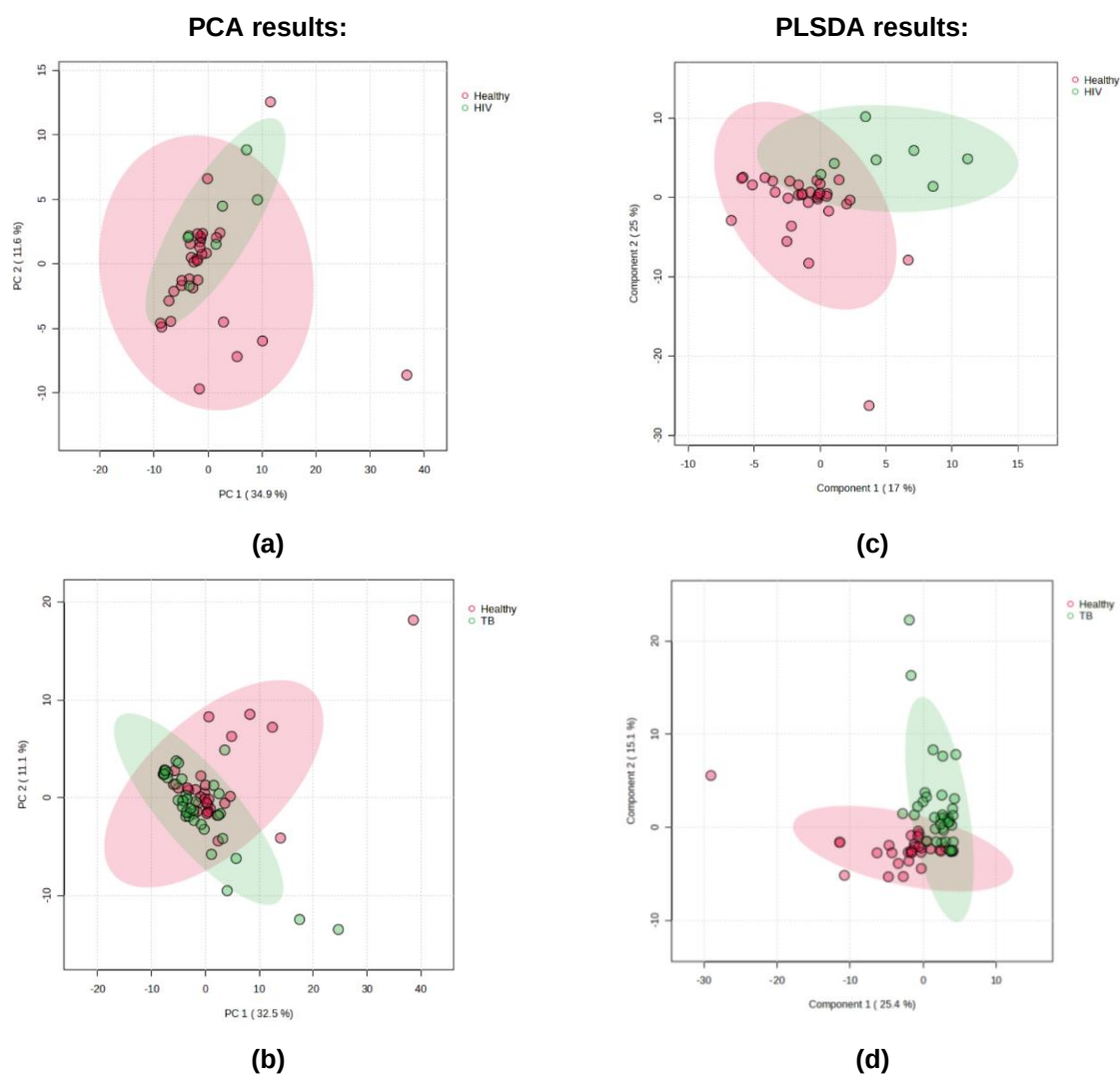


Figure 4.11: Comparative Analysis Using PCA and PLS-DA. This figure illustrates the group comparisons using PCA for inherent data structure visualisation, and PLS-DA for group classification to emphasise the discriminative features between the groups. Specifically, panel (a) shows the PCA of HIV-positive individuals compared to healthy controls, highlighting differences in their metabolic profiles. Panel (b) presents the PCA for TB-positive individuals versus healthy controls. Panel (c) illustrates the PLS-DA for HIV-positive individuals compared to healthy controls, emphasising discriminative features between these groups. Finally, panel (d) shows the PLS-DA for TB-positive individuals versus healthy controls.

4.4.4.2. Cross-Validation and Permutation Testing

Given that PLS-DA is prone to overfitting data, cross-validation was performed to evaluate the model's reliability and predictive power. Cross-validation in this context involves assessing how well the PLS-DA model generalises to an independent dataset. A key metric in this process is the Q2 statistic, which estimates the predictive capacity of the model. Generally, a Q2 value closer to 1 indicates a model with good predictive ability. In our study, the cross-validation results, depicted in Figure 4.12, showed Q2 values below 0.8, which implies some limitations in the model's predictive accuracy.

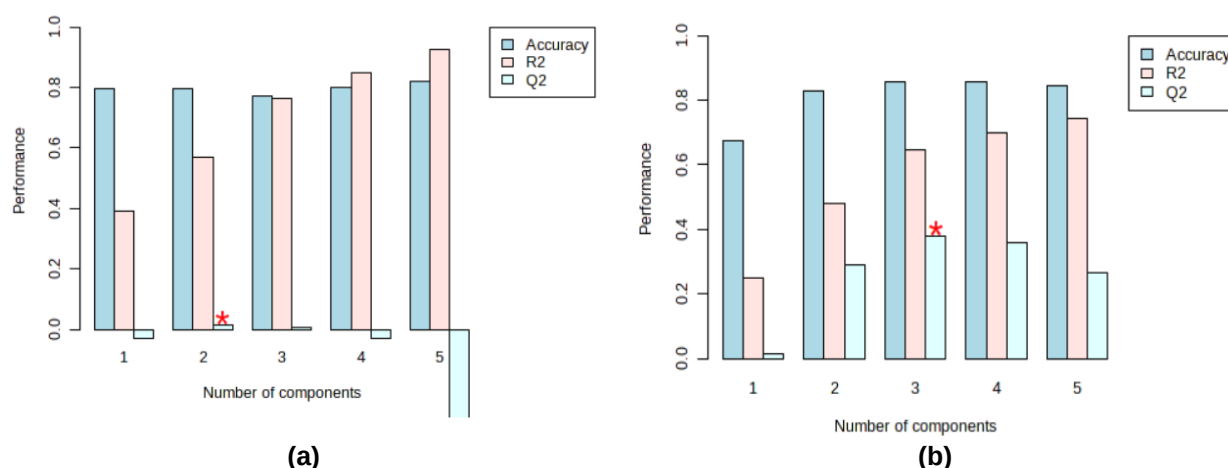


Figure 4.12: Cross-validation test results. Here, (a) displays the cross-validation tests for the HIV-positive group versus healthy controls, and (b) shows the TB-positive group versus healthy controls.

Permutation testing, performed in addition to cross-validation as depicted in Figure 4.13, further assessed the statistical reliability of the PLS-DA models. This method evaluates the significance of the test statistics under the strong null hypothesis by comparing shuffled statistics to actual statistics, offering a robust means to discern the reliability of the models against random variations (Rice & Lumley, 2008).

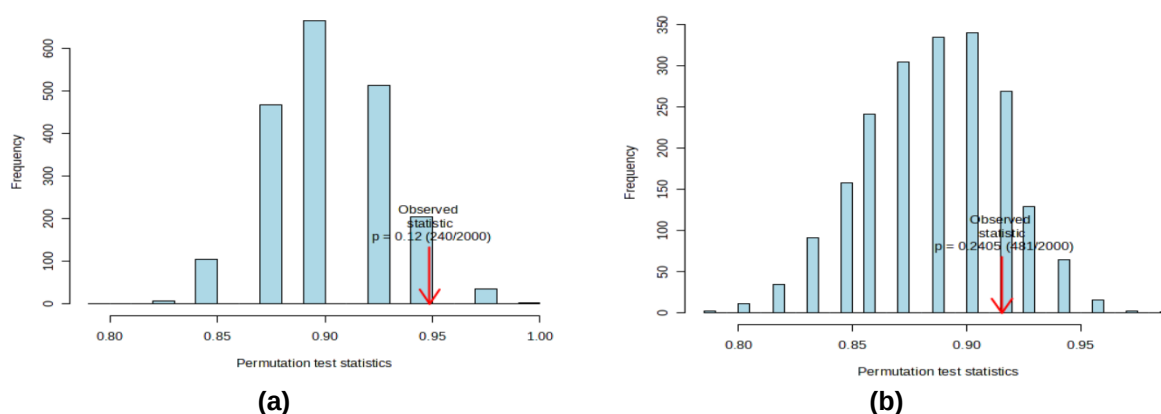


Figure 4.13: Permutation test results. Permutation tests results for the (a) HIV-positive group versus healthy controls and the (b) TB-positive group versus healthy controls.

In our study, the permutation test for the HIV-positive group compared to the healthy controls [Figure 4.13(a)] resulted in a p-value of 0.12 (240/2000), indicating a poor permutation test percentage. Similarly, the comparison of the TB-positive group versus the healthy controls [Figure 4.13(b)] yielded a p-value of 0.2405 (481/2000), also suggesting a poor predictive capacity. These results imply that the observed group differences might not be statistically significant. One potential reason for these limited results could be the small cohort size, as a larger sample size is generally required for a reliable estimation of empirical p-values. Due to these considerations, PLS-DA was cautiously interpreted and not relied upon for identifying significant metabolite markers in this study.

These tests collectively highlight the importance of rigorous validation methods in multivariate analysis, especially when dealing with complex data sets as in this study. The results underscore the need for cautious interpretation of the PLS-DA model outcomes.

4.5. DISCUSSION

The outcomes of this study present a nuanced picture of the utility of PCA and PLS-DA in this study. Due to the limitations revealed by the cross-validation and permutation tests, neither PCA nor PLS-DA could confidently be used for further analysis. These results underscore the necessity for more robust analytical methods or larger datasets in future research endeavours to yield conclusive findings.

In line with the study's exploratory nature, the focus was directed toward the metabolite markers identified through both the Kruskal-Wallis (n=15; Table 4.3) and Wilcoxon signed-rank (n=22; Table 4.4) tests. This comprehensive list (n=25, Table 4.5) was considered for an in-depth analysis to explore their potential roles and implications in the context of HIV, TB, and their co-infection. Thus, the discussion aims to bridge the observed metabolic alterations with their biological significance, thereby contributing to the existing body of knowledge on these diseases and offering a nuanced understanding of their metabolic effects.

4.5.1. Amino Acids and Related Metabolites

Diseases profoundly influence amino acid metabolism, integral to numerous metabolic processes. Understanding these changes offers insights into the complexities of metabolic pathways and the impact of diseases like HIV and TB. In this study, examining the amino acid profiles revealed distinct alterations associated with HIV, TB, and co-infection, offering a window into the possible metabolic disturbances caused by these conditions.

In the HIV-positive group, compared to healthy controls, there was an observed trend in glycine levels which may reflect complex changes in protein biosynthesis and breakdown associated

with HIV infection, although this did not reach high statistical significance in our analysis. As the simplest amino acid and a significant neurotransmitter, glycine plays a role in various physiological processes, including the synthesis of DNA, phospholipids, and collagen. While an increase in glycine during HIV infection might be linked to its known anti-inflammatory properties (Aguayo-Cerón *et al.*, 2023), this observation should be interpreted cautiously. Given that HIV infection leads to a heightened immune response and systemic inflammation, a potential increase in glycine levels could rise as part of the body's natural response to mitigate this inflammation. Furthermore, glycine is integral to the body's antioxidant defences, which may become increasingly active in response to the elevated oxidative stress typically observed in HIV infection. The virus' impact on liver function, pivotal in amino acid metabolism, could also be a contributing factor in this observed trend, but further investigation is needed. Additionally, the nutritional status of individuals with HIV, affecting the absorption and metabolism of various nutrients including amino acids, could play a role in the observed increase in glycine (Zhang *et al.*, 2023).

In HIV-positive individuals, the observed increased trends in pentanedioic acid and hydracrylic acid points towards broader alterations in both amino acid and fatty acid metabolism due to the infection. Pentanedioic acid, a dicarboxylic acid, arises not only from amino acid degradation but also plays a role in the metabolism of lysine and tryptophan (Gholson *et al.*, 1959), indicating its involvement in various metabolic pathways. Similarly, hydracrylic acid, a breakdown product of branched-chain amino acids and propanoic acid (Wishart *et al.*, 2022), may increase due to changes in fatty acid metabolism and gut microbiota composition associated with HIV (González-Hernández *et al.*, 2019), however with the low statistical significance achieved in our analysis further investigation is needed to fully confirm the role of hydracrylic acid as a significant metabolite between these disease states. These metabolic shifts can be attributed to the virus' impact on cellular processes and the body's adaptive response to the infection, encompassing both amino acid utilisation and lipid metabolism.

In TB-positive patients, the elevated levels of 3-phenyllactic acid can be directly linked to cachexia, a condition characterised by severe muscle wasting (Luies & Du Preez, 2020a). This condition triggers increased protein breakdown, releasing amino acids like phenylalanine into the bloodstream. The resulting accumulation of phenylalanine, due to decreased activity of phenylalanine hydroxylase, leads to its conversion into phenylpyruvic acid and subsequently into phenyllactic acid (Amalia *et al.*, 2022; Luies & Loots, 2016). This chain of metabolic events highlights the profound impact of TB on amino acid metabolism, where the systemic inflammation and immune activation characteristic of TB contribute to metabolic stress, affecting enzymes involved in amino acid processing. The occurrence of cachexia is further substantiated by the elevated levels of *tyrosine* in the TB group. Tyrosine, whose systemic concentration depends on protein consumption and release from tissue reservoirs, can

increase due to cachexia-induced protein breakdown and altered insulin production impacting amino acid metabolism (Luies & Loots, 2016).

The increase in quinolinic acid in the TB-positive and HIV/TB co-infected groups, as observed in this study, is indicative of the enhanced and chronic inflammatory response characteristic of both HIV and TB. Quinolinic acid, as part of the kynurenine pathway, is not only a marker of inflammation but also plays a role in neurodegenerative processes (Blumenthal *et al.*, 2012; Luies & Loots, 2016). Its neurotoxic effects can contribute to neurological symptoms in diseases like TB, especially when the central nervous system is involved. Since its levels were increased the most in the co-infected group, it implies that the synergistic impact of these infections results in a more pronounced inflammatory response, further activating the kynurenine pathway and leading to increased quinolinic acid production. Elevated quinolinic acid levels in these patient groups could thus signal not only heightened inflammation but also potential neurological damage or dysfunction, further supporting its potential as a biomarker for neurological involvement in TB and HIV/TB co-infection.

4.5.2. Carbohydrates, TCA cycle, and Energy Metabolism

Diseases significantly impact the tricarboxylic acid (TCA) cycle and overall energy metabolism, which are central to cellular energy production and various metabolic processes. Infections like HIV and TB can cause profound alterations in these pathways, reflecting the extensive metabolic demands and the body's adaptive responses to the disease state.

The elevated average levels of citric acid in HIV-positive individuals when compared across all comparative groups (Appendix D), signify the extensive metabolic adaptations occurring in response to HIV infection. Citric acid, a vital component of the TCA cycle, is crucial for energy production. In HIV-infected cells, the virus-induced inflammation and macrophage activation lead to increased citric acid production and mitochondrial export, essential for meeting the heightened energy demands of these cells (Infantino *et al.*, 2011; Koppensteiner *et al.*, 2012; Patil *et al.*, 2019). This metabolic shift is a critical adaptation of the immune system to counteract the virus, underlining the profound impact of HIV on cellular energy metabolism and the TCA cycle.

When comparing TB-positive individuals to healthy controls, a notable decrease in certain carbohydrates, including L-threonic acid, ribitol, and meso-erythritol, was observed. L-Threonic acid, a metabolite of ascorbic acid (vitamin C), is crucial for antioxidant defence and collagen synthesis (Bergamini *et al.*, 2008; Lloyd *et al.*, 1987). The observed decrease in L-threonic acid might signify a deficiency in vitamin C. This deficiency could stem from inadequate dietary intake, a common issue in the socioeconomically challenged areas where this cohort reside. Additionally, TB's impact on the body, such as decreased appetite and potential malabsorption,

can lead to a further reduction in vitamin C intake and absorption. Furthermore, the chronic inflammation and immune activation characteristic of TB can induce oxidative stress, which might deplete vitamin C and its metabolites, including L-threonic acid, as they are utilised more rapidly in immune responses and tissue repair processes.

Ribitol, a sugar alcohol derived from the reduction of ribose, is a crucial component of riboflavin (vitamin B2) and flavin mononucleotide, playing a vital role in the pentose phosphate pathway. In TB patients, a decrease in ribitol levels might be attributed to TB's effect on glucose metabolism and the subsequent disruption of the pentose phosphate pathway. This metabolic imbalance contributes to the observed decrease in ribitol (Badolia et al., 2020). Similarly, gluconolactone, an oxidised derivative of glucose, is reduced in TB patients, reflecting the impact of TB on glucose metabolism. The highest concentrations of gluconolactone were observed in healthy controls, suggesting an unaltered glucose metabolism compared to TB patients. Moreover, TB-associated malnutrition and weight loss hinder effective nutrient metabolism, affecting compounds like ribitol and gluconolactone (Gupta et al., 2009). The inflammation and immune cell activation characteristic of TB alter metabolic demands and processes, including those involving ribitol. Oxidative stress, another common feature of TB, disrupts metabolic functions and further influences ribitol metabolism. It is also plausible that *Mtb* directly or indirectly affects ribitol metabolism, altering synthesis or degradation pathways (Shastri et al., 2018).

Meso-erythritol, a sugar alcohol, demonstrates contrasting levels in TB and HIV patients, underscoring the distinct metabolic influences of these diseases. In TB patients, a decrease in meso-erythritol aligns with impaired glucose tolerance, a common diabetic condition also seen in TB (Luies & Loots, 2016). TB's metabolic disturbances, coupled with nutritional challenges like malnutrition and weight loss, lead to reduced levels of meso-erythritol. The inflammatory immune response in TB can further alter metabolic processes, reducing meso-erythritol synthesis or availability. Cachexia in TB, leading to reduced insulin secretion, may also explain the metabolic alterations observed in TB patients, including the relationship with diabetes (Luies & Loots, 2016). In contrast, the HIV-positive group exhibited the highest meso-erythritol levels. This could be due to HIV's impact on glucose metabolism, potentially enhancing meso-erythritol synthesis. Additionally, the oxidative stress induced by HIV may elevate meso-erythritol levels as part of the body's antioxidant response, reflecting a compensatory mechanism to counteract the infection's oxidative burden (Chiriac et al., 2022).

4.5.3. Dicarboxylic Acids and Lipids

Dicarboxylic acids and lipids play vital roles in cellular metabolism, including energy production, signal transduction, and maintaining cell structural integrity. Alterations in these

compounds can be indicative of significant metabolic changes occurring in disease states like HIV and TB.

In TB-positive patients, a noticeable decrease in octanoic acid was observed compared to healthy controls. Octanoic acid is involved in energy metabolism and hormone regulation. This decline may be due to nutritional deficits and metabolic stress prevalent in TB. Disruption in hormones such as ghrelin, which is acylated with octanoic acid, might contribute to its reduced levels (Chang et al., 2013; Yurt et al., 2013). Additionally, TB's impact on glucose tolerance and energy metabolism could lead to decreased octanoic acid synthesis or increased utilisation.

Similarly, propanoic acid levels were reduced in TB patients, likely reflecting metabolic impairments and nutritional deficiencies common in the disease. These changes can influence gut microbiota, subsequently altering short-chain fatty acid production, including propanoic acid (Ranjbar et al., 2021). Conversely, as mentioned previously, propanoic acid was found to be elevated in HIV patients, which could be attributed to HIV-induced gut dysbiosis. These changes impact fatty acid metabolism, leading to increased production of propanoic acid (González-Hernández et al., 2019).

2-Methylene-butane-1,4-diol showed increased levels in both HIV and TB groups compared to controls. This elevation might be linked to the diseases' heightened immune response and inflammation, affecting lipid metabolism. Alterations in lipid mobilisation and synthesis due to infection-related metabolic demands, along with potential medication effects, could be contributing factors (Funderburg & Mehta, 2016).

In the HIV-positive group, there was an increase in methylmalonic acid, which is associated with a B12 deficiency (Boutin et al., 2020; Boutin et al., 2022), a condition that can arise due to HIV's impact on nutrient absorption and metabolism (Tang & Smit, 1998). Further investigation is required to fully understand the relationship between methylmalonic acid and B12 deficiency in HIV patients.

In summary, HIV infection might be associated with an increase in certain dicarboxylic acids and lipid metabolites, potentially due to an altered amino acid metabolism, nutrient deficiencies, and effects on gut microbiota. In contrast, TB infection seems to lead to a decrease in these metabolites, reflecting the disease's impact on nutritional status, hormonal regulation, and overall metabolic disruption. These findings highlight the significant metabolic shifts in HIV and TB, emphasizing the need to understand these changes for effective therapeutic interventions.

4.5.4. Gut Microbiota Imbalance

The interplay between TB and the gut microbiota is a critical aspect of understanding TB's comprehensive impact on human health. Beyond the primary respiratory system infection, TB significantly alters the gut microbiome, which is closely linked to the host immune system (Luies & Du Preez, 2020a). A study by Winglee *et al.* (2014) on *Mtb* infection in humans and animal models revealed marked changes in gut microbial composition, including reduced diversity and alterations in bacterial populations linked to inflammation. Maji *et al.* (2018) reported similar findings, highlighting distinct microbiome profiles in TB patients, characterised by decreased abundance of key bacterial genera. These alterations in gut microbiota, particularly the reduced production of short-chain fatty acids (SCFAs) like the previously mentioned propanoic acid, play a vital role in gut integrity and immune modulation. The reduction in SCFAs can exacerbate inflammation and contribute to the cachexia often observed in TB patients. Recent metabolomic studies have brought attention to altered phenylalanine metabolism in TB patients, suggesting this change is a consequence of dysregulated gut microbiota activity (Das *et al.*, 2015; Luies & Loots, 2016). This alteration has also been correlated with poor treatment outcomes (Luies *et al.*, 2017). It is important to recognise that variations in the gut microbiota related to TB might also be influenced by external factors such as diet, body mass, and environmental conditions (Das *et al.*, 2015).

In this study, TB-positive patients exhibited a notable reduction in gut microbiota-associated markers, specifically hippuric acid and valeric acid, compared to healthy controls. Interestingly, these markers were most elevated in the healthy control group, when compared to all other diseased groups, indicating a clear disruption in gut microbiota balance due to *Mtb* infection. Hippuric acid, primarily derived from the degradation of plant polyphenols and aromatic amino acids like phenylalanine and tryptophan, is a product of gut microbial activity (Li *et al.*, 2008). It is formed through the conjugation of benzoic acid with glycine and is subsequently excreted in urine (Wishart *et al.*, 2022). A decrease in hippuric acid excretion can indicate liver dysfunction (Kohlstaedt & Helmer, 1936), aligning with the urine dipstick results from our study (Section 4.4.2). The complex relationship of TB with the liver, manifesting as diffuse hepatic involvement, granulomatous hepatitis, or localised tuberculoma or abscess (Sonika & Kar, 2012), might also explain the reduced levels of hippuric acid in TB patients. This liver involvement in TB may also contribute to the poor treatment outcomes observed in many cases (De Villiers & Loots, 2013).

Valeric acid, another gut microbiota-derived marker, is a short-chain fatty acid produced from undigested carbohydrates (Skrzypecki *et al.*, 2020; Tan *et al.*, 2014). Its beneficial role for the host is well documented. The observed decrease in valeric acid in TB patients suggests an alteration in the gut microbiota during the host's immune response against TB (Liu *et al.*, 2021).

This finding is in line with existing research that highlights TB's influence on the host's entire metabolome, including the normal functioning of gut microbiota (Luies & Du Preez, 2020a).

In contrast, 2,3-butanediol and 2,3-dihydroxybutanoic acid were elevated in HIV-positive individuals compared to healthy controls, reflecting HIV's impact on gut microbiota. 2,3-Butanediol is produced through bacterial fermentation of pyruvate, a process influenced by HIV-related gut microbiome alterations. HIV infection disrupts the gut lining, increasing intestinal permeability and altering the microbial environment. This disruption promotes the growth of acetoin-producing microorganisms, where acetoin is a precursor to 2,3-butanediol, leading to its increased urinary levels (Mills & Walker, 1989). Enhanced carbohydrate fermentation due to altered gut microbiota in HIV also contributes to this increase. Therefore, elevated 2,3-butanediol in HIV patients signifies not just the direct effect of the virus on the gut but broader metabolic changes. On the other hand, 2,3-dihydroxybutanoic acid is a sugar acid and can be a product of bacterial metabolism. Its increased levels in HIV-positive individuals may be linked to an altered gut microbial balance, leading to an overgrowth of certain bacteria that produce 2,3-dihydroxybutanoic acid. Additionally, HIV infection is associated with various metabolic disturbances, including altered carbohydrate metabolism, which could also contribute to the increase in this compound. Given this compound's bacterial origins, its elevation in HIV could be indicative of a shift in gut microbial populations or an altered metabolic state due to the infection (Dillon *et al.*, 2016).

These findings collectively underscore the significant role of the gut microbiota in disease pathogenesis and its potential implications in patient outcomes and novel therapeutic strategies that target gut microbiome restoration as part of the treatment plan.

4.5.5. DNA Damage and Oxidative Stress

Both HIV and TB are closely associated with increased oxidative stress and DNA damage, significantly impacting their pathogenesis and progression. In the case of HIV, this association primarily stems from the HIV-induced activation of immune cells, leading to reactive oxygen species (ROS) production and oxidative stress (Westreich *et al.*, 2017). Oxidative stress, characterised by an imbalance between ROS production and the body's detoxification capabilities, can lead to cellular damage, including DNA damage, lipid peroxidation, and protein modifications. The integration of HIV into the host genome exacerbates DNA damage, disrupting normal cellular processes and contributing to the genomic instability of infected cells (Ivanov *et al.*, 2016).

This study's findings, including the elevation of 2-deoxyribolactone in the HIV-positive group compared to healthy controls, align with these mechanisms. 2-Deoxyribolactone, a product of deoxyribose oxidation in DNA during oxidative stress (Chan *et al.*, 2010), is indicative of the

DNA damage induced by HIV. Specifically, the HIV viral protein R is known to directly cause DNA damage, stalling DNA replication, and leading to both double and single-stranded DNA breaks (Li et al., 2020), thereby increasing 2-deoxyribolactone formation. The complex relationship between HIV infection, oxidative stress, and DNA damage not only plays a role in HIV pathogenesis but also affects ART efficacy and overall health outcomes in HIV-infected individuals (Ivanov *et al.*, 2016). Therefore, addressing oxidative stress and its resultant DNA damage is crucial in the comprehensive management of HIV.

4.5.6. Dietary Origin

The findings of elevated levels of 2-hydroxycyclohexane-carboxylic acid and diaveridine, which are not typically produced endogenously, likely originate from dietary sources or environmental exposures. Both metabolites were found to be elevated in the TB patients compared to controls. 2-Hydroxy-cyclohexane-carboxylic acid, a member of the cyclohexanols class, is often found in various foods and natural products (Wishart *et al.*, 2022). It is important to note that our control group included individuals from the same township as the TB patients, but also comprised individuals from surrounding non-township areas. This diversity in the control group's socio-economic backgrounds could account for differences in dietary habits and environmental exposures. Consequently, the higher intake of certain foods or traditional medicines containing cyclohexanols in the TB cohort could explain the increased levels of this metabolite.

Diaveridine, an antibacterial agent, is often used in animal produce to inhibit folic acid metabolism in bacteria. Apart from being significantly increased in the TB cohort, the highest levels were detected in the co-infected group when compared to all other groups. This suggests exposure through diet, potentially through the consumption of animal products treated with diaveridine. It seems that the metabolic systems of the co-infected individuals might handle the excretion and processing of diaveridine differently due to the compounded effects of both HIV and TB.

This observation points to the importance of considering dietary habits and environmental exposures in understanding the metabolic profiles of patients with infectious diseases like TB and HIV. It also underscores the potential impact of socio-economic and cultural factors on disease progression and treatment outcomes in these populations.

4.5.7. Exogenous Compounds

Lastly, ephedrine, a medication commonly used as a decongestant and bronchodilator to alleviate breathing problems (Limberger et al., 2013), was significantly increase in TB patients compared to healthy controls. Since TB is a respiratory disease, the use of ephedrine or

ephedrine-like medications can be used to manage respiratory symptoms. Although participants were advised to indicate any medication consumed, this might have been omitted, and therefore these results should be interpreted with caution.

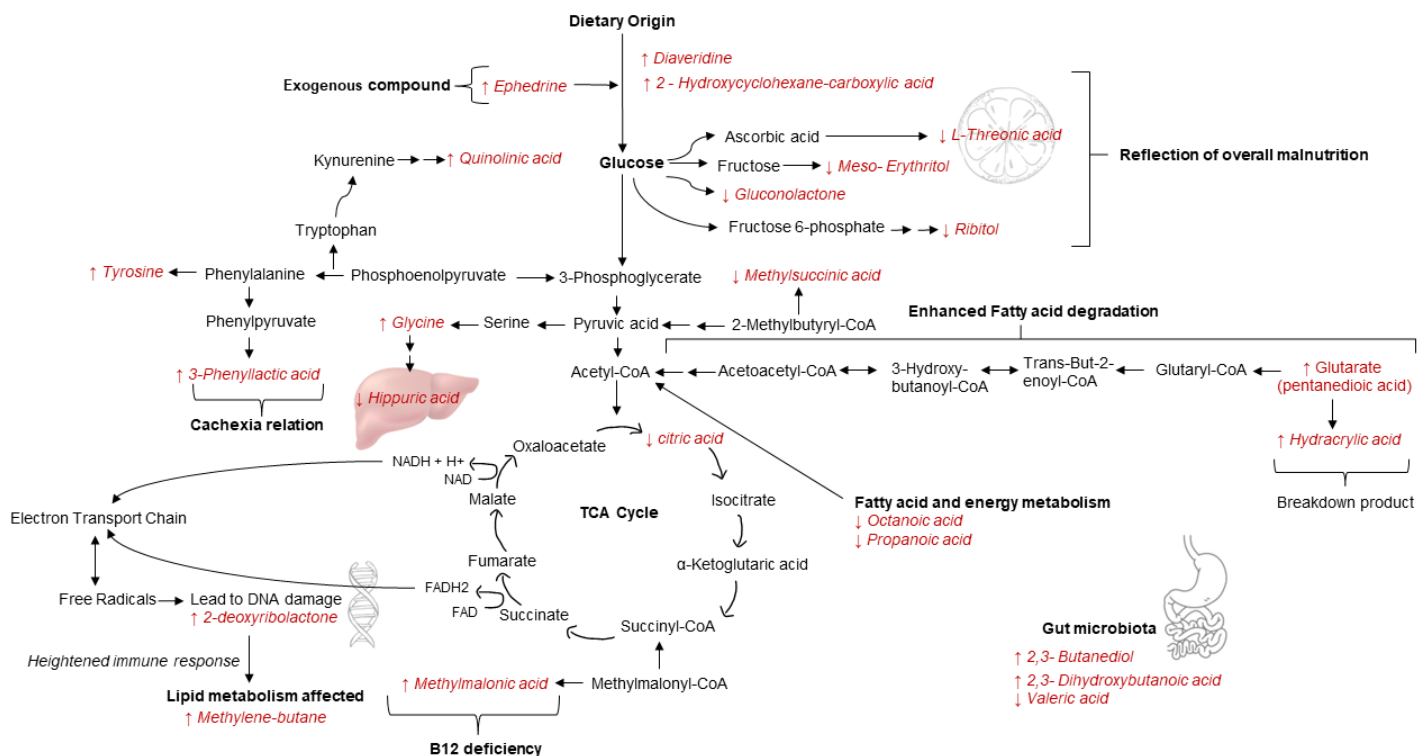


Figure 4.14: Metabolic alterations in TB, HIV, and HIV/TB co-infection. This highlights key metabolic pathways impacted, including amino acid, fatty acid, and carbohydrate metabolism. The interplay between TB and HIV leads to distinct changes in various metabolites, reflecting the diseases' profound effects on the body's metabolic processes. Abbreviations: *Mtb*: *Mycobacterium tuberculosis*; *HIV*: human immunodeficiency virus; *TCA*: tricarboxylic acid; *NADH*: nicotinamide adenine dinucleotide (reduced); *FADH₂*: flavin adenine dinucleotide (reduced); *CoA*: coenzyme A; *NAD⁺*: nicotinamide adenine dinucleotide (oxidized); *FAD*: flavin adenine dinucleotide (oxidised).

4.6. CONCLUSION

This explorative study reveals a complex interplay between HIV, TB, and co-infection with their respective metabolic profiles, significantly influenced by both diseases. The utilisation of Kruskal-Wallis and Wilcoxon signed-rank tests highlighted key metabolic alterations, indicating profound impacts on amino acid metabolism, energy pathways, gut microbiota, and oxidative stress responses. Specifically, HIV infection's possible association with altered amino acid degradation, heightened oxidative stress, and gut microbiota disruptions. Elevated levels of compounds like glycine, pentanedioic acid, hydracrylic acid, and 2,3-butanediol in HIV-positive individuals underscore these metabolic shifts.

In TB patients, the comprehensive metabolic alterations include cachexia, indicated by altered amino acid metabolism, and a marked impact on nutrient absorption and metabolic processes. The findings highlight TB's influence on systemic inflammation, immune activation, and its interconnection with diabetes. This is evidenced by the impaired glucose metabolism and compromised insulin secretion, showcasing TB's extensive metabolic implications, enhancing our understanding of its far-reaching effects.

Co-infection demonstrates a compounded metabolic impact, particularly evident in the elevated levels of quinolinic acid and diaveridine. These findings suggest a synergistic effect of HIV and TB on inflammatory responses and potentially on the processing of dietary and exogenous compounds.

4.7. REFERENCES

- Aguayo-Cerón, K.A., Sánchez-Muñoz, F., Gutierrez-Rojas, R.A., Acevedo-Villavicencio, L.N., Flores-Zarate, A.V., Huang, F., ... Romero-Nava, R. 2023. Glycine: The Smallest Anti-Inflammatory Micronutrient. *International Journal of Molecular Sciences*, 24(14):11236.
- Amalia, F., Syamsunarno, M.R.A., Triatin, R.D., Fatimah, S.N., Chaidir, L. & Achmad, T.H. 2022. The Role of Amino Acids in Tuberculosis Infection: A Literature Review. *Metabolites*, 12(10):933.
- Badolia, R., Ramadurai, D.K., Abel, E.D., Ferrin, P., Taleb, I., Shankar, T.S., ... Yin, M. 2020. The role of nonglycolytic glucose metabolism in myocardial recovery upon mechanical unloading and circulatory support in chronic heart failure. *Circulation*, 142(3):259-274.
- Badri, M., Ehrlich, R., Wood, R., Pulerwitz, T. & Maartens, G. 2001. Association between tuberculosis and HIV disease progression in a high tuberculosis prevalence area. *The International Journal of Tuberculosis and Lung Disease*, 5(3):225-232.
- Baker, M., Das, D., Venugopal, K. & Howden-Chapman, P. 2008. Tuberculosis associated with household crowding in a developed country. *Journal of Epidemiology & Community Health*, 62(8):715-721.
- Bergamini, P., Marchesi, E., Marchi, A., Bertolasi, V., Fogagnolo, M. & Canella, A. 2008. Formation of l-Threonic Acid from l-Ascorbic Acid Oxidative Ring Opening and its Coordination to PtII: X-ray Crystal Structures of [Pt (threonato-O, O')(PPh₃)₂] and [Pt (oxalato)(PPh₃)₂]. Wiley Online Library.
- Bi, H., Guo, Z., Jia, X., Liu, H., Ma, L. & Xue, L. 2020. The key points in the pre-analytical procedures of blood and urine samples in metabolomics studies. *Metabolomics*, 16:1-15.
- Blumenthal, A., Nagalingam, G., Huch, J.H., Walker, L., Guillemin, G.J., Smythe, G.A., ... Saunders, B.M. 2012. M. tuberculosis induces potent activation of IDO-1, but this is not essential for the immunological control of infection. *PloS one*, 7(5):e37314.
- Boehme, C.C., Nabeta, P., Hillemann, D., Nicol, M.P., Shenai, S., Krapp, F., ... Rustomjee, R. 2010. Rapid molecular detection of tuberculosis and rifampin resistance. *New England Journal of Medicine*, 363(11):1005-1015.
- Boshoff, H.I. & Barry, C.E. 2005. Tuberculosis—metabolism and respiration in the absence of growth. *Nature Reviews Microbiology*, 3(1):70-80.
- Bouatra, S., Aziat, F., Mandal, R., Guo, A.C., Wilson, M.R., Knox, C., ... Liu, P. 2013. The human urine metabolome. *PloS one*, 8(9):e73076.
- Boutin, M., Presse, N., Martineau, T., Perreault, A., Gaudreau, P. & Auray-Blais, C. 2020. Mass spectrometry analysis of urinary methylmalonic acid to screen for metabolic vitamin B12 deficiency in older adults. *Bioanalysis*, 12(10):693-705.

- Boutin, M., Presse, N., Allard, D., Martineau, T., Gaudreau, P. & Auray-Blais, C. 2022. Methylmalonic acid analysis using urine filter paper samples to screen for metabolic vitamin B12 deficiency in older adults. *Bioanalysis*, 14(9):615-626.
- Boyd, J. & Barratt, J. 2011. Interpretation and management of abnormal dipstick urinalysis. *Medicine*, 39(6):312-316.
- Brennan, P.J. 2003. Structure, function, and biogenesis of the cell wall of *Mycobacterium tuberculosis*. *Tuberculosis*, 83(1-3):91-97.
- Broadhurst, D., Goodacre, R., Reinke, S.N., Kuligowski, J., Wilson, I.D., Lewis, M.R. & Dunn, W.B. 2018. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics*, 14:1-17.
- Brookmeyer, R. 1991. Reconstruction and future trends of the AIDS epidemic in the United States. *Science*, 253(5015):37-42.
- Bruchfeld, J., Correia-Neves, M. & Källénus, G. 2015. Tuberculosis and HIV coinfection. *Cold Spring Harbor perspectives in medicine*, 5(7):a017871.
- Chakaya, J., Khan, M., Ntoumi, F., Aklillu, E., Fatima, R., Mwaba, P., ... Katoto, P.D. 2021. Global Tuberculosis Report 2020—Reflections on the Global TB burden, treatment and prevention efforts. *International journal of infectious diseases*, 113:S7-S12.
- Chan, W., Chen, B., Wang, L., Taghizadeh, K., Demott, M.S. & Dedon, P.C. 2010. Quantification of the 2-deoxyribonolactone and nucleoside 5'-aldehyde products of 2-deoxyribose oxidation in DNA and cells by isotope-dilution gas chromatography mass spectrometry: Differential effects of γ -radiation and Fe²⁺–EDTA. *Journal of the American Chemical Society*, 132(17):6145-6153.
- Chang, S.W., Pan, W.S., Lozano Beltran, D., Oleyda Baldeomar, L., Solano, M.A., Tuero, I., ... Gilman, R.H. 2013. Gut hormones, appetite suppression and cachexia in patients with pulmonary TB. *PloS one*, 8(1):e54564.
- Chiriac, A.P., Ghilan, A., Serban, A.-M., Macsim, A.-M., Barga, A., Doroftei, F., ... Sandu, A.-I. 2022. Preparation of an Antioxidant Assembly Based on a Copolymacrolactone Structure and Erythritol following an Eco-Friendly Strategy. *Antioxidants*, 11(12):2471.
- Coetzee, D., Hildebrand, K., Boule, A., Maartens, G., Louis, F., Labatula, V., ... Goemaere, E. 2004. Outcomes after two years of providing antiretroviral treatment in Khayelitsha, South Africa. *Aids*, 18(6):887-895.
- Cohen, M.S., Chen, Y.Q., McCauley, M., Gamble, T., Hosseinipour, M.C., Kumarasamy, N., ... Pilotto, J.H. 2016. Antiretroviral therapy for the prevention of HIV-1 transmission. *New England Journal of Medicine*, 375(9):830-839.
- Das, M.K., Bishwal, S.C., Das, A., Dabral, D., Badireddy, V.K., Pandit, B., ... Nanda, R.K. 2015. Deregulated tyrosine–phenylalanine metabolism in pulmonary tuberculosis patients. *Journal of proteome research*, 14(4):1947-1956. <http://pubs.acs.org/doi/pdfplus/10.1021/acs.jproteome.5b00016>
- De Villiers, L. & Loots, D.T. 2013. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Current metabolomics*, 1(4):306-317.
- DeSanto, D., Velen, K., Lessells, R., Makgopa, S., Gumede, D., Fielding, K., ... Chetty-Makkan, C.M. 2023. A qualitative exploration into the presence of TB stigmatization across three districts in South Africa. *BMC Public Health*, 23(1):1-10.
- Dillon, S., Lee, E., Kotter, C., Austin, G., Gianella, S., Siewe, B., ... Robertson, C. 2016. Gut dendritic cell activation links an altered colonic microbiome to mucosal and systemic T-cell activation in untreated HIV-1 infection. *Mucosal immunology*, 9(1):24-37.
- Dorward, J., Khubone, T., Gate, K., Ngobese, H., Sookrajh, Y., Mkhize, S., ... Baisley, K. 2021. The impact of the COVID-19 lockdown on HIV care in 65 South African primary care clinics: an interrupted time series analysis. *The Lancet HIV*, 8(3):e158-e165.
- Eriksson, L., Byrne, T., Johansson, E., Trygg, J. & Vikström, C. 2013. *Multi-and megavariable data analysis basic principles and applications*. 1. Umetrics Academy.
- Foster, N., Vassall, A., Cleary, S., Cunnam, L., Churchyard, G. & Sinanovic, E. 2015. The economic burden of TB diagnosis and treatment in South Africa. *Social science & medicine*, 130:42-50.

- Funderburg, N.T. & Mehta, N.N. 2016. Lipid abnormalities and inflammation in HIV infection. *Current HIV/AIDS Reports*, 13:218-225.
- Garcia, M., Luecke, E., Mayo, A.J., Scheckter, R., Ndase, P., Kiweewa, F.M., ... Singh, N. 2021. Impact and experience of participant engagement activities in supporting dapivirine ring use among participants enrolled in the phase III MTN-020/ASPIRE study. *BMC Public Health*, 21:1-11.
- Gholson, R., Sanders, D. & Henderson, L. 1959. Glutaric acid: A product of tryptophan metabolism. *Biochemical and Biophysical Research Communications*, 1(2):98-100.
- Gomez, C. & Hope, T.J. 2005. The ins and outs of HIV replication. *Cellular microbiology*, 7(5):621-626.
- González-Hernández, L.A., Ruiz-Briseño, M.d.R., Sánchez-Reyes, K., Alvarez-Zavala, M., Vega-Magaña, N., López-Iñiguez, A., ... Ramos-Solano, M. 2019. Alterations in bacterial communities, SCFA and biomarkers in an elderly HIV-positive and HIV-negative population in western Mexico. *BMC Infectious Diseases*, 19:1-12.
- Grzegorzewicz, A.E., De Sousa-D'Auria, C., McNeil, M.R., Huc-Claustre, E., Jones, V., Petit, C., ... Belardinelli, J.M. 2016. Assembling of the Mycobacterium tuberculosis cell wall core. *Journal of Biological Chemistry*, 291(36):18867-18879.
- Gupta, K.B., Gupta, R., Atreja, A., Verma, M. & Vishvkarma, S. 2009. Tuberculosis and nutrition. *Lung India*, 26(1):9-16.
- Harris, R.C., Quaife, M., Weerasuriya, C., Gomez, G.B., Sumner, T., Bozzani, F. & White, R.G. 2022. Cost-effectiveness of routine adolescent vaccination with an M72/AS01E-like tuberculosis vaccine in South Africa and India. *Nature Communications*, 13(1):602.
- Hecke, T.V. 2012. Power study of anova versus Kruskal-Wallis test. *Journal of Statistics and Management Systems*, 15(2-3):241-247.
- Herbert, C., Luies, L., Loots, D.T. & Williams, A.A. 2023. The metabolic consequences of HIV/TB co-infection. *BMC Infectious Diseases*, 23(1):536.
- Hernandez-Vargas, E.A. & Middleton, R.H. 2013. Modeling the three stages in HIV infection. *Journal of theoretical biology*, 320:33-40.
- Hirsch, M.S. & Curran, J. 1990. *Human immunodeficiency viruses. Biology and medical aspects*. 4. New York, NY: Lippincott-Raven Publishers. (Virology).
- HIV/AIDS, J.U.N.P.o. 2021. *UNAIDS Data 2021*. Available from: https://www.unaids.org/en/resources/documents/2021/2021_unaids_data Date of access: 2021.
- Hollander, M., Wolfe, D.A. & Chicken, E. 2013. *Nonparametric statistical methods*. John Wiley & Sons.
- Holmberg, S. & Rothstein, B. 2011. Dying of corruption. *Health Economics, Policy and Law*, 6(4):529-547.
- Horsburgh Jr, C.R. 2004. Priorities for the treatment of latent tuberculosis infection in the United States. *New England journal of medicine*, 350(20):2060-2067.
- Hoskin, T. 2012. Parametric and nonparametric: Demystifying the terms. In. Mayo Clinic. pp. 1-5.
- Hussey, G. 2007. Prospects for new tuberculosis vaccines. *South African Medical Journal*, 97(10):1001-1002.
- Infantino, V., Convertini, P., Cucci, L., Panaro, M.A., Di Noia, M.A., Calvello, R., ... Iacobazzi, V. 2011. The mitochondrial citrate carrier: a new player in inflammation. *Biochemical Journal*, 438(3):433-436.
- Ioerger, T.R., Koo, S., No, E.-G., Chen, X., Larsen, M.H., Jacobs Jr, W.R., ... Sacchettini, J.C. 2009. Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa. *PloS one*, 4(11):e7778.
- Ivanov, A.V., Valuev-Elliston, V.T., Ivanova, O.N., Kochetkov, S.N., Starodubova, E.S., Bartosch, B. & Isaguliant, M.G. 2016. Oxidative stress during HIV infection: mechanisms and consequences. *Oxidative medicine and cellular longevity*, 2016,
- Jensen, B.-E.O., Knops, E., Cords, L., Lübke, N., Salgado, M., Busman-Sahay, K., ... Wittner, M. 2023. In-depth virological and immunological characterization of HIV-1 cure after CCR5Δ32/Δ32 allogeneic hematopoietic stem cell transplantation. *Nature Medicine*:1-5.

- Karim, S.S.A. & Baxter, C. 2010. Microbicides & their implications in HIV prevention. *The Indian Journal of Medical Research*, 132(6):656.
- Knechel, N.A. 2009. Tuberculosis: pathophysiology, clinical features, and diagnosis. *Critical care nurse*, 29(2):34-43.
- Kohlstaedt, K. & Helmer, O.M. 1936. A study of the hippuric acid excretion as a test of hepatic function. *The American Journal of Digestive Diseases*, 3:459-466.
- Koppensteiner, H., Brack-Werner, R. & Schindler, M. 2012. Macrophages and their relevance in Human Immunodeficiency Virus Type I infection. *Retrovirology*, 9:1-11.
- Kwan, C.K. & Ernst, J.D. 2011. HIV and tuberculosis: a deadly human syndemic. *Clinical microbiology reviews*, 24(2):351-376.
- Laher, F., Bekker, L.-G., Garrett, N., Lazarus, E.M. & Gray, G.E. 2020. Review of preventative HIV vaccine clinical trials in South Africa. *Archives of virology*, 165:2439-2452.
- Lawn, S.D. & Nicol, M.P. 2011. Xpert® MTB/RIF assay: development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. *Future microbiology*, 6(9):1067-1082.
- Levy, J.A. 1993. Pathogenesis of human immunodeficiency virus infection. *Microbiological reviews*, 57(1):183-289.
- Li, D., Lopez, A., Sandoval, C., Nichols Doyle, R. & Fregoso, O.I. 2020. HIV Vpr modulates the host DNA damage response at two independent steps to damage DNA and repress double-strand DNA break repair. *Mbio*, 11(4):e00940-00920.
- Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., ... Zhang, M. 2008. Symbiotic gut microbes modulate human metabolic phenotypes. *Proceedings of the National Academy of Sciences*, 105(6):2117-2122.
- Liebenberg, C., Luies, L. & Williams, A.A. 2021. Metabolomics as a Tool to Investigate HIV/TB Co-Infection. *Frontiers in Molecular Biosciences*, 8:1-19.
- Lienhardt, C. 2001. From exposure to disease: the role of environmental factors in susceptibility to and development of tuberculosis. *Epidemiologic reviews*, 23(2):288-301.
- Limberger, R.P., Jacques, A.L.B., Schmitt, G.C. & Arbo, M.D. 2013. Pharmacological Effects of Ephedrine 38.
- Liu, Y., Wang, J. & Wu, C. 2021. Microbiota and tuberculosis: a potential role of probiotics, and postbiotics. *Frontiers in Nutrition*, 8:626254.
- Lloyd, L.L., Warner, F., White, C. & Kennedy, J. 1987. Quantitative reversed phase HPLC analysis of L-ascorbic acid (vitamin C) and identification of its degradation products. *Chromatographia*, 24:371-376.
- LoBue, P.A. & Catanzaro, A. 1997. The diagnosis of tuberculosis. *Disease-a-Month*, 43(4):185-246.
- Lönnroth, K., Jaramillo, E., Williams, B.G., Dye, C. & Raviglione, M. 2009. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Social science & medicine*, 68(12):2240-2246.
- Loveday, M. & Zweigenthal, V. 2011. TB and HIV integration: obstacles and possible solutions to implementation in South Africa. *Tropical Medicine & International Health*, 16(4):431-438.
- Luies, L. & Loots, D.T. 2016. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics*, 12(3):1-9.
- Luies, L. & Du Preez, I. 2020a. The echo of pulmonary tuberculosis: mechanisms of clinical symptoms and other disease-induced systemic complications. *Clinical Microbiology Reviews*, 33(4):e00036-00020.
- Luies, L. & Du Preez, I. 2020b. The echo of pulmonary tuberculosis: mechanisms of clinical symptoms and other disease-induced systemic complications. *Clinical Microbiology Reviews*, 33(4):1-19.
- Luies, L., Van Reenen, M., Ronacher, K., Walzl, G. & Loots, D.T. 2017. Predicting tuberculosis treatment outcome using metabolomics. *Biomarkers in medicine*, 11(12):1057-1067. <https://doi.org/10.2217/bmm-2017-0133>
- Mahamed, D., Boule, M., Ganga, Y., Mc Arthur, C., Skroch, S., Oom, L., ... Rampersad, S. 2017. Intracellular growth of Mycobacterium tuberculosis after macrophage cell death leads to serial killing of host cells. *Elife*, 6:e22028.

- Maji, A., Misra, R., Dhakan, D.B., Gupta, V., Mahato, N.K., Saxena, R., ... Singh, A. 2018. Gut microbiome contributes to impairment of immunity in pulmonary tuberculosis patients by alteration of butyrate and propionate producers. *Environmental microbiology*, 20(1):402-419.
- Mansoor, L.E., Lewis, L., Naicker, C.L., Harkoo, I., Dawood, H., Naidoo, K., ... Abdool Karim, S.S. 2022. Prospective study of oral pre-exposure prophylaxis initiation and adherence among young women in KwaZulu-Natal, South Africa. *Journal of the International AIDS Society*, 25(7):e25957.
- Mantzaris, E. & Pillay, P. 2021. Legal profession and corruption in health care: some reflective realities in South Africa. *Frontiers in Public Health*, 9:688049.
- Maskew, M., Bor, J., MacLeod, W., Carmona, S., Sherman, G.G. & Fox, M.P. 2019. Adolescent HIV treatment in South Africa's national HIV programme: a retrospective cohort study. *The lancet HIV*, 6(11):e760-e768.
- Mawar, N., Sahay, S., Pandit, A. & Mahajan, U. 2005. The third phase of HIV pandemic: social consequences of HIV/AIDS stigma & discrimination & future needs. *Indian Journal of Medical Research*, 122(6):471.
- McKight, P.E. & Najab, J. 2010. Kruskal-wallis test. *The corsini encyclopedia of psychology*:1-1.
- Mdluli, P. & Dunga, S. 2022. Determinants of poverty in South Africa using the 2018 general household survey data. *Journal of Poverty*, 26(3):197-213.
- Miller, C.M., Ketlhapile, M., Rybasack-Smith, H. & Rosen, S. 2010. Why are antiretroviral treatment patients lost to follow-up? A qualitative study from South Africa. *Tropical Medicine & International Health*, 15:48-54.
- Mills, G. & Walker, V. 1989. Urinary excretion of 2, 3-butanediol and acetoin by babies on a special care unit. *Clinica chimica acta*, 179(1):51-59.
- Mlangeni, P. 2023. South Africa one of 30 high TB-burdened countries. <https://www.sabcnews.com/sabcnews/south-africa-one-of-30-high-tb-burdened-countries/> Date of access:
- Montagnier, L. & Clavel, F. 1994. *Human immunodeficiency viruses*. 2. London: London: Academic Press Ltd. (Encyclopedia of virology.).
- Moosa, A., Gengiah, T.N., Lewis, L. & Naidoo, K. 2019. Long-term adherence to antiretroviral therapy in a South African adult patient cohort: a retrospective study. *BMC infectious diseases*, 19(1):1-12.
- Nicol, M.P., Workman, L., Isaacs, W., Munro, J., Black, F., Eley, B., ... 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. *The Lancet infectious diseases*, 11(11):819-824.
- Nyamweya, S., Hegedus, A., Jaye, A., Rowland-Jones, S., Flanagan, K.L. & Macallan, D.C. 2013. Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis. *Reviews in medical virology*, 23(4):221-240.
- Olivier, C., Allen, B. & Luies, L. 2023. Optimising a urinary extraction method for non-targeted GC-MS metabolomics. *Scientific Reports*, 13(1):17591.
- Olivier, I. & Loots, D. 2011. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play. *Journal of Cell and Tissue Research*, 11(1):2655-2671.
- Organization, W.H. 2022. *WHO announces landmark changes in treatment of drug-resistant tuberculosis*. <https://www.who.int/news/item/15-12-2022-who-announces-landmark-changes-in-treatment-of-drug-resistant-tuberculosis> Date of access:
- Pang, Z., Chong, J., Zhou, G., de Lima Morais, D.A., Chang, L., Barrette, M., ... Xia, J. 2021. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic acids research*, 49(W1):W388-W396.
- Patil, N.K., Bohannon, J.K., Hernandez, A., Patil, T.K. & Sherwood, E.R. 2019. Regulation of leukocyte function by citric acid cycle intermediates. *Journal of leukocyte biology*, 106(1):105-117.
- Paton, N., Ng, Y., Chee, C., Persaud, C. & Jackson, A. 2003. Effects of tuberculosis and HIV infection on whole-body protein metabolism during feeding, measured by the [15N] glycine method. *American journal of clinical nutrition*, 78:319-325. <https://dx.doi.org/10.1093/ajcn/78.2.319> 10.1093/ajcn/78.2.319
- Paton, N.I. & Ng, Y.-M. 2006. Body composition studies in patients with wasting associated with tuberculosis. *Nutrition*, 22(3):245-251. <https://dx.doi.org/10.1016/j.nut.2005.06.009> 10.1016/j.nut.2005.06.009

- Pelissari, D.M. & Diaz-Quijano, F.A. 2017. Household crowding as a potential mediator of socioeconomic determinants of tuberculosis incidence in Brazil. *PLoS One*, 12(4):e0176116.
- Peltzer, K., Parker, W., Mabaso, M., Makonko, E., Zuma, K. & Ramlagan, S. 2012. Impact of national HIV and AIDS communication campaigns in South Africa to reduce HIV risk behaviour. *The Scientific World Journal*, 2012,
- Ranjbar, R., Vahdati, S.N., Tavakoli, S., Khodaie, R. & Behboudi, H. 2021. Immunomodulatory roles of microbiota-derived short-chain fatty acids in bacterial infections. *Biomedicine & Pharmacotherapy*, 141:111817.
- Roberts, J.R. 2007. Urine dipstick testing: everything you need to know. *Emergency Medicine News*, 29(6):24-27.
- Russell, D.G. 2001. Mycobacterium tuberculosis: here today, and here tomorrow. *Nature reviews Molecular cell biology*, 2(8):569-578.
- Ryan, D., Robards, K., Prenzler, P.D. & Kendall, M. 2011. Recent and potential developments in the analysis of urine: a review. *Analytica chimica acta*, 684(1-2):17-29.
- Saccenti, E., Hoefsloot, H.C., Smilde, A.K., Westerhuis, J.A. & Hendriks, M.M. 2014. Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics*, 10:361-374.
- Salazar-Austin, N., Dowdy, D.W., Chaisson, R.E. & Golub, J.E. 2019. Seventy years of tuberculosis prevention: efficacy, effectiveness, toxicity, durability, and duration. *American journal of epidemiology*, 188(12):2078-2085.
- Schneider, H., Blaauw, D., Gilson, L., Chabikuli, N. & Goudge, J. 2006. Health systems and access to antiretroviral drugs for HIV in Southern Africa: service delivery and human resources challenges. *Reproductive health matters*, 14(27):12-23.
- Schoeman, J.C., Du Preez, I. & Loots, D.T. 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(2):301-311.
- Sharp, P.M. & Hahn, B.H. 2011. Origins of HIV and the AIDS pandemic. *Cold Spring Harbor perspectives in medicine*, 1(1):a006841.
- Shastri, M.D., Shukla, S.D., Chong, W.C., Dua, K., Peterson, G.M., Patel, R.P., ... O'Toole, R.F. 2018. Role of oxidative stress in the pathology and management of human tuberculosis. *Oxidative medicine and cellular longevity*, 2018,
- Shaw, G.M. & Hunter, E. 2012. HIV transmission. *Cold Spring Harbor perspectives in medicine*, 2(11):a006965.
- Sierra, S., Kupfer, B. & Kaiser, R. 2005. Basics of the virology of HIV-1 and its replication. *Journal of clinical virology*, 34(4):233-244.
- Singer, M. 2006. A dose of drugs, a touch of violence, a case of AIDS, part 2: Further conceptualizing the SAVA syndemic. *Free Inquiry in Creative Sociology*, 34(1):39-54.
- Skrzypecki, J., Niewęłowska, K. & Samborowska, E. 2020. Valeric acid, a gut microbiota product, penetrates to the eye and lowers intraocular pressure in rats. *Nutrients*, 12(2):387.
- Smith, I. 2003. Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clinical microbiology reviews*, 16(3):463-496.
- Smuts, I., van der Westhuizen, F.H., Louw, R., Mienie, L.J., Engelke, U.F., Wevers, R.A., ... Reinecke, C.J. 2013. Disclosure of a putative biosignature for respiratory chain disorders through a metabolomics approach. *Metabolomics*, 9:379-391.
- Sonika, U. & Kar, P. 2012. Tuberculosis and liver disease: management issues. *Tropical Gastroenterology*, 33(2):102-106.
- Stoneburner, R.L., Chiasson, M.A., Weisfuse, I.B. & Thomas, P.A. 1990. The epidemic of AIDS and HIV-1 infection among heterosexuals in New York City. *AIDS*, 4(2):99-106.
- Tan, J., McKenzie, C., Potamitis, M., Thorburn, A.N., Mackay, C.R. & Macia, L. 2014. The role of short-chain fatty acids in health and disease. *Advances in immunology*, 121:91-119.
- Tang, A.M. & Smit, E. 1998. Selected vitamins in HIV infection: a review. *AIDS patient care and STDs*, 12(4):263-273.

- Taylor, S. 2023. Nonparametric Tests. <https://corporatefinanceinstitute.com/resources/data-science/nonparametric-tests/>
- Tiberi, S., Carvalho, A.C.C., Sulis, G., Vaghela, D., Rendon, A., Mello, F.C.d.Q., ... Pontali, E. 2017. The cursed duet today: Tuberculosis and HIV-coinfection. *La Presse Médicale*, 46(2):e23-e39.
- Tucker, T.J. & Mazithulela, G. 2004. Development of an AIDS vaccine: perspective from the South African AIDS Vaccine Initiative. *BMJ*, 329(7463):454-456.
- Van Lettow, M., Fawzi, W.W. & Semba, R.D. 2003. Triple trouble: The role of malnutrition in tuberculosis and human immunodeficiency virus co-infection. *Nutrition reviews*, 61(3):81-90. 10.131/nr.2003.marr.81-90
- Venkatesan, P. 2023. HIV vaccine trial failure. *The Lancet Infectious Diseases*, 23(4):410.
- Visca, D., Ong, C., Tiberi, S., Centis, R., D'ambrosio, L., Chen, B., ... Dalcolmo, M. 2021. Tuberculosis and COVID-19 interaction: a review of biological, clinical and public health effects. *Pulmonology*, 27(2):151-165.
- Wani, R.L.S. 2013. Tuberculosis 2: Pathophysiology and microbiology of pulmonary tuberculosis. *South Sudan Medical Journal*, 6(1):10-12.
- Warner, D.F. 2015. Mycobacterium tuberculosis metabolism. *Cold Spring Harbor perspectives in medicine*, 5(4):a021121.
- Weinrib, R., Browne, E., Shapley-Quinn, M., van der Straten, A., Beksinska, M., Mgodhi, N., ... Ju, S. 2020. Perspectives from young south African and Zimbabwean women on attributes of four (placebo) vaginal microbicide delivery forms. *AIDS and Behavior*, 24:637-647.
- Westreich, D., Cates, J., Cohen, M., Weber, K.M., Seidman, D., Cropsey, K., ... Mehta, C.C. 2017. Smoking, HIV, and risk of pregnancy loss. *AIDS (London, England)*, 31(4):553.
- Winglee, K., Eloie-Fadrosch, E., Gupta, S., Guo, H., Fraser, C. & Bishai, W. 2014. Aerosol *Mycobacterium tuberculosis* infection causes rapid loss of diversity in gut microbiota. *PLOS one*, 9(5):e97048.
- Wishart, D.S., Guo, A., Oler, E., Wang, F., Anjum, A., Peters, H., ... Lee, B.L. 2022. HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Research*, 50(D1):D622-D631.
- World Health Organization. 2021. *Global Tuberculosis Report*. Geneva: World Health Organization. Available from: <https://www.who.int/publications/i/item/9789240037021> Date of access: 2021.
- Yurt, S., Erman, H., Korkmaz, G., Kosar, A., Uysal, P., Gelisgen, R., ... Uzun, H. 2013. The role of feed regulating peptides on weight loss in patients with pulmonary tuberculosis. *Clinical biochemistry*, 46(1-2):40-44.
- Zhang, J., Chen, Y., Wang, M., Zhong, L., Li, L., Yuan, Z. & Zou, S. 2023. Amino acid metabolism dysregulation associated with inflammation and insulin resistance in HIV-infected individuals with metabolic disorders. *Amino Acids*:1-11.

CHAPTER 5: CONCLUSION AND FUTURE PROSPECTS

Since a comprehensive discussion was already given above in the results sections, the current chapter will only provide a summary on the main findings in the context of the aim and objectives. Future prospects will also be discussed.

5.1. FINAL CONCLUSIONS

Comprehensive method development is essential for yielding optimized and reliable results. In Section A Part 1, we determined the optimal urine volume for our in-house organic acid (OA) extraction, focusing on analytical repeatability, metabolome coverage, and metabolite recovery. The full-micro method demonstrated high repeatability and metabolome coverage, while the semi-micro method excelled in metabolite recovery. However, the full-micro method faced practical challenges, in handling small volumes, such as the precise transfer of organic phases. The small volume of Na₂SO₄ used in the full-micro method also presented difficulties during decanting, necessitating the use of a glass Pasteur pipette for accurate transfer, unlike the free-hand decanting in other methods. Given these challenges, the semi-micro method is recommended for GC-MS applications due to its balance of practicality and performance. The full-micro method, despite its limitations, shows promise with potential future automation and method validation improvements.

Hereafter (Part 2), we assessed five urine sample preparation methods for non-targeted GC-MS metabolomics applications, again evaluating repeatability, metabolome coverage, and metabolite recovery. A low-volume DA method outperformed the others, ranking first in all validation parameters and identifying 91 unique peaks across multiple compound classes with high metabolite recovery. This method was then further optimised (Part 3) by investigating an additional drying step between two-step derivatization and the use of urease sample pre-treatment. Findings showed that while an additional drying step was beneficial, the method did not benefit from urease pre-treatment. Therefore, urease pretreatment is recommended only for specific analyses targeting urea-related measurements, in which case sonication, instead of heating, is the preferred method to catalyse this reaction.

Following the optimisation of a urine extraction method, it was applied to patient samples to reveal the metabolic profile characterising HIV/TB co-infection (Section B). Using Kruskal-Wallis and Wilcoxon signed-rank tests, key metabolic alterations were identified, indicating significant impacts on amino acid metabolism, energy pathways, gut microbiota, and oxidative stress responses. In HIV individuals, metabolic shifts included potentially altered amino acid degradation, heightened oxidative stress, and gut microbiota disruptions, underscored by elevated levels of compounds like glycine, pentanedioic acid, hydracrylic acid, and

2,3-butanediol. In TB individuals, altered amino acid metabolism and nutrient absorption impacts were observed, indicating a state of cachexia and highlighting TB's influence on systemic inflammation, immune activation, and interconnection with diabetes, evidenced by impaired glucose metabolism and compromised insulin secretion. In co-infection, a compound metabolic effect was noted, particularly highlighted by elevated levels of quinolinic acid and diaveridine, suggesting a synergistic effect of HIV and TB on inflammatory responses and the processing of dietary and exogenous compounds.

As a hypothesis-generating study, this research laid a foundational framework for future targeted GC-MS metabolomics studies aimed at further elucidating the mechanisms underlying the HIV/TB syndemic and their potentiation. However, it is crucial to acknowledge the limitations of this study to inform future improvements. One significant limitation was the sample size, which was not large enough to provide statistically robust conclusions. This constrained the depth of statistical analysis possible, thereby limiting the interpretive power of the results obtained. Another challenge faced was the financial constraints in acquiring standards, which restricted our ability to achieve a higher MSI identification level for metabolic peaks. As a result, we had to rely on in-house spectral libraries, which might not provide the same level of confidence as external, standardised references. While this study successfully serves as an exploratory investigation to aid future co-infection diagnostic studies, its generalisability is constrained by these methodological and resource limitations. Future studies would benefit from a larger sample size and more comprehensive metabolite identification to provide more definitive insights into the metabolic interplay in HIV/TB co-infection.

5.2. FUTURE PROSPECTS

In light of the results obtained, future prospects could include:

- Comprehensive metabolic analysis in HIV/TB cohorts and comparative studies: Utilise data from this study to distinguish metabolic consequences of TB from HIV cohorts for improved diagnosis and treatment of co-infection. Also, compare these findings with metabolic profiles from other infectious diseases for unique metabolic signatures identification.
- Development of diagnostic models using metabolomic data: Explore the use of identified metabolites to construct diagnostic models for HIV/TB co-infection and review existing literature on the metabolomic impacts of TB and HIV for novel diagnostic insights.
- Immunometabolism, targeted metabolomics, and amino acid analysis: Investigate the immunometabolic profiles in TB infection, conduct targeted metabolomics studies focusing on specific pathways or metabolites, and perform detailed targeted amino acid analysis.

- Advanced analytical techniques for focused analyses, including NMR for glucose analysis: Incorporate techniques like NMR and high-resolution mass spectrometry for improved metabolite identification, with a special focus on carbohydrates such as glucose, considering the TB-diabetes link.
- Larger cohort studies in varied geographical regions: Conduct studies with larger cohorts across different regions to account for variations in *Mtb* strains and achieve more robust characterisation of metabolic profiles.
- Data analysis using machine learning and integrative approaches: Apply machine learning for complex data analysis and integrate metabolomics data with genomic and proteomic analyses for a holistic understanding of disease mechanisms.
- Population-based biomarker validation and immune response studies: Validate potential biomarkers in diverse populations and investigate how metabolic changes due to HIV and TB affect the immune system's response.

APPENDIX A

Olivier, C., Allen, B. and Luies, L. (2023). Optimising a urinary extraction method for non-targeted GC–MS metabolomics. *Scientific Reports*, 13(1), p.17591.
(<https://doi.org/10.1038/s41598-023-44690-7>)



OPEN

Optimising a urinary extraction method for non-targeted GC–MS metabolomics

Cara Olivier¹ , Bianca Allen¹ & Laneke Luies¹ ✉

Urine is ideal for non-targeted metabolomics, providing valuable insights into normal and pathological cellular processes. Optimal extraction is critical since non-targeted metabolomics aims to analyse various compound classes. Here, we optimised a low-volume urine preparation procedure for non-targeted GC–MS. Five extraction methods (four organic acid [OA] extraction variations and a “direct analysis” [DA] approach) were assessed based on repeatability, metabolome coverage, and metabolite recovery. The DA method exhibited superior repeatability, and achieved the highest metabolome coverage, detecting 91 unique metabolites from multiple compound classes comparatively. Conversely, OA methods may not be suitable for all non-targeted metabolomics applications due to their bias toward a specific compound class. In accordance, the OA methods demonstrated limitations, with lower compound recovery and a higher percentage of undetected compounds. The DA method was further improved by incorporating an additional drying step between two-step derivatization but did not benefit from urease sample pre-treatment. Overall, this study establishes an improved low-volume urine preparation approach for future non-targeted urine metabolomics applications using GC–MS. Our findings contribute to advancing the field of metabolomics and enable efficient, comprehensive analysis of urinary metabolites, which could facilitate more accurate disease diagnosis or biomarker discovery.

Metabolomics refers to the quantitative measurement of dynamic metabolic changes in a system responding to genetic modifications or physiological stimuli, either external (e.g., drugs) or internal (e.g., nutrients)¹. It involves the analysis of the metabolome, which can be defined as the complete set of small molecules (metabolites) present in a biological sample. Multi-level profiling of metabolites can be conducted in an unbiased manner through the comparison of sets of biological samples².

Gas chromatography mass spectrometry (GC–MS) is considered the “gold standard” for metabolomics studies³ due to its many advantages, including its suitability for both volatile and non-volatile compound analyses (after derivatization), affordability in terms of low running costs, high sensitivity with excellent chromatographic and mass resolution, good dynamic range, and compound identification using mass spectral library matching (of which extensive commercial libraries are available). Furthermore, GC–MS provides additional and orthogonal data (such as retention time), reproducible chromatographic separations (making results comparable between different laboratories and analysts), and the ability to distinguish stereoisomers. GC–MS also has shorter run times with lower bleed⁴.

Although GC–MS has some disadvantages, such as the co-elution of compounds and slower scan rates, most of these can be compensated for, for example, with two-dimensional coupling (GCxGC) or by coupling the system to time-of-flight mass spectrometry (TOFMS), respectively⁵. These approaches broaden the application of GC–MS substantially⁶. However, since GC–MS is limited to volatile, thermally stable compounds—and most natural occurring compounds in a biological matrix are non-volatile and unstable at high temperatures—samples need to be derivatized prior to analysis⁷. Derivatization refers to the process of chemically modifying a compound, exploiting its polarity, and boiling point, including the interaction with a GC’s capillary column to achieve compound separation. Hence, most methods require two-step derivatization, employing oximation followed by silylation⁸, although the latter is often used as a one-step derivatization approach. These derivatization methods will be further explored in this publication.

Urine has proven to be an excellent sample matrix for non-targeted metabolomics studies because it reflects both normal and pathological cellular processes, providing a holistic assessment of metabolic profiles. Urine is a particularly valuable sample matrix due to its ease of collection, straightforward handling and processing

Human Metabolomics, North-West University, Potchefstroom Campus, Private Bag X6001, Box 269, Potchefstroom 2520, NW, South Africa. ✉email: laneke.luies@nwu.ac.za

in the laboratory, and the reduced risk of nosocomial transmission to healthcare and laboratory personnel. These advantages make urine a promising alternative biological sample, especially given the challenges in the collection and yield of other biological samples in certain diseases⁹. Early therapy intervention and improved treatment effectiveness depend on the insights derived from urine metabolomics. Therefore, urine metabolomics can be employed to rule out potential metabolic defects, assess nutritional issues, and detect signs of infections or overgrowth (including pathogen or flora). To maximise the utility of urinary metabolome data for disease characterisation and treatment strategies, effective sample preparation, coupled with selective extractions tailored to the chosen analytical technique, becomes critically important¹⁰.

For urine GC–MS analyses, organic acid (OA) metabolite extraction remains a popular option. The principle of an OA extraction is typically based on a two-step liquid–liquid extraction (LLE; also known as partitioning)¹¹ which involves the use of two immiscible solvents¹². However, LLE is better suited for targeted investigations in clinical settings, such as diagnosing organic acidurias⁴. Optimal extraction is critical since non-targeted metabolomics aims to analyse various compound classes at varying concentrations. To this end, recent advances in technology have led to the development of more efficient and cost-effective methods for compound extraction, such as the direct analysis (DA) method, for non-targeted metabolomics.

The DA method involves deproteinisation, sample concentration, and derivatization¹³. Although straightforward, various studies suggest an additional drying step between two-step derivatization^{14,15} to improve reproducibility (since residual water and other contaminants can be removed), reduce matrix effects (by removing water and other impurities that can interfere with ionisation and compound detection), and increase sensitivity (since background noise is reduced while the signal-to-noise ratio is increased).

One significant challenge in non-targeted metabolomics using urine samples, is the presence of urea, which can obscure other metabolites in the sample with the same retention time and/or interfere with the derivatization process, leading to incomplete chemical transformation and the formation of urea-derived artifacts^{7,8}. To address this issue, various studies suggest using the enzyme urease to remove excess urea from urine samples prior to GC–MS analysis^{16,17}. This also has the advantage of avoiding column overloading, peak distortions, lower chances of co-eluting peaks, and lower coefficient of variation (CV) values compared to urease-non-treated samples^{16,18}. However, the use of urease pre-treatment has its disadvantages, including the occurrence of unwanted chemical transformations and secondary enzymatic reactions that can alter the metabolic profile^{19,20}, and the inability to detect several compounds in urease-treated samples compared to non-treated samples¹⁶. As a result, the use of urease pre-treatment in non-targeted metabolomics remains a subject of debate in the scientific community.

In this investigation, we aimed to optimise a low-volume urine preparation procedure for non-targeted GC–MS analyses, covering various compound classes. In Part A, five methods were evaluated based on repeatability, metabolome coverage, and metabolite recovery. An OA extraction method (Method 1) was compared with modified versions, namely the use of an additional extraction solvent (Method 2), two-step derivatization (Method 3), and a combination of Methods 2 and 3 (Method 4). A DA method (Method 5) was also tested. In Part B, the superior method was further optimised by investigating if an additional drying step between two-step derivatization (Method 6) and urease sample pre-treatment (Method 7) is indeed beneficial for disease characterisation studies.

Methods

Chemicals

The following reagents were acquired from Sigma Aldrich (St. Louis, Missouri, USA) and Merck (Darmstadt, Germany): 3-phenylbutyric acid, methoxyamine hydrochloride (MOX-HCl), anhydrous sodium sulphate (Na₂SO₄), urease, glucose, 4-aminobutyric acid (GABA), L-alanine, L-leucine, L-phenylalanine, L-tryptophan, L-tyrosine, ascorbic acid, citric acid, succinic acid, palmitic acid, arabinose, ribose, N,O-bis(trimethylsilyl)trifluoroacetamide (BSTFA) with 1% trimethylsilyl chloride (TMCS), and pyridine. Other chemicals used were acetonitrile, ethyl acetate, hydrogen chloride and diethyl ether from Burdick and Jackson brand (Honeywell International Inc., Muskegon, USA).

Standards

A spiking mixture/solution (50 ppm) was prepared prior to the experiments, using known amounts of compounds from different compound classes. This mixture was used to determine the percentage recovery of the known compounds introduced to the samples. Three different compound classes were included: amino acids (GABA, L-alanine, L-leucine, L-phenylalanine, L-tryptophan, L-tyrosine), organic acids (ascorbic acid, citric acid, succinic acid), and a fatty acid (palmitic acid).

An internal standard (IS; 50 ppm) solution containing 3-phenylbutyric acid was also prepared separately.

Sample preparation and ethical approval

The investigation was done according to the Declaration of Helsinki and International Conference of Harmonization guidelines. Ethical approval was obtained from the Ethics Committee of the North-West University for a larger study (NWU-00355-20-A1), as well as for this sub-study (NWU-00355-20-A1-03). All recruited participants gave written informed consent.

A single quality control (QC) sample, compiled from ethically approved, previously collected healthy urine samples (n = 32) from which multiple smaller (100 µL) aliquots were made, was used in this experiment. Before an aliquot was transferred, the material was thoroughly mixed by vortex. The sample aliquots were thawed at room temperature, and to eliminate any crystals and make sure that any precipitated material has dissolved, the urine was centrifuged at 15 700×g for 5 min at room temperature.

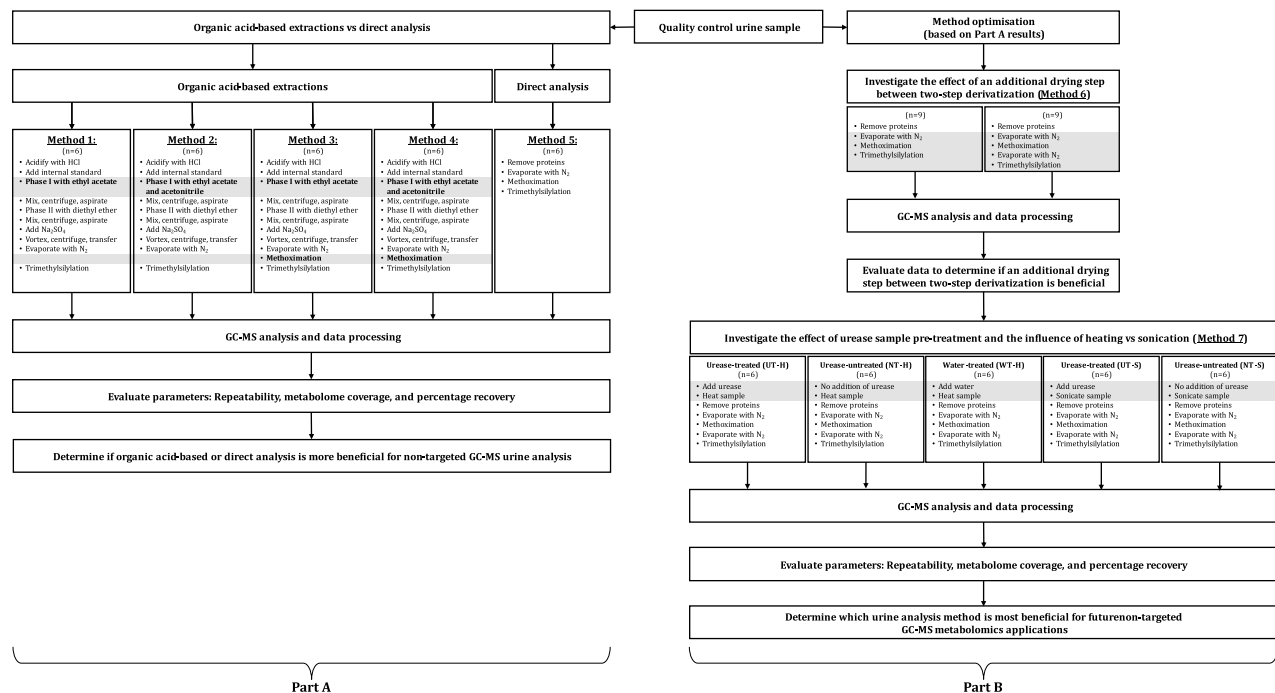


Figure 1. Summary of the experimental design. This study comprises a Part A (to evaluate five extraction methods) and Part B (to optimise a urine preparation procedure for non-targeted GC-MS analyses). The variations between the methods are indicated in grey sections. For Methods 1–5, six QC aliquots were used each, while Method 6 required 18 aliquots and Methods 7 needed 30 aliquots.

Metabolite extraction

An illustration of the experimental design is available in Fig. 1. For Part A of this investigation, five methods were evaluated based on repeatability, metabolome coverage, and metabolite recovery. An OA extraction method (previously optimised for a reduced urine volume; Method 1) was compared with modified versions of this method, namely the use of acetonitrile as an additional extraction solvent (Method 2), two-step derivatization to include both oximation and silylation (Method 3), as well as a combination of Methods 2 and 3 (Method 4). These methods were also compared to a DA method (Method 5).

In Part B, the superior method was further optimised by investigating the effect(s) of an additional drying step between two-step derivatization (Method 6). Hereafter, urease sample pre-treatment was evaluated (Method 7) by comparing urease-treated samples to both urease-non-treated and water-treated samples. Furthermore, the effect of the heating process that urease pre-treatment requires was investigated by comparing the urease-treated to urease-non-treated samples subjected to both heating and sonication.

Methods 1–5 was conducted as a single experiment, for which six QC aliquots were used each ($n = 30$), while Method 6 was conducted in a second experiment, requiring 18 aliquots, and Method 7 as a third experiment, which needed 30 aliquots.

Method 1 (traditional OA extraction)²¹

Six drops of 5N HCl were added to 100 μ L urine samples filled up to 500 μ L with a 9% saline solution ($n = 6$), followed by the addition of IS (100 μ L, final concentration of 50 ppm). Hereafter, half of the samples ($n = 3$) were spiked (referred to as the “pre-spike samples”), and ethylacetate (6 mL) was added to all samples. The samples were mixed (30 min) and centrifuged (845xg for 5 min at 4 $^{\circ}$ C), and the organic phase was collected. Diethylether (3 mL) was added to the remaining aqueous phase, and the samples were again mixed, centrifuged, and the organic phase collected and combined with the previously collected phase. Approximately 3 g Na₂SO₄ was added to each sample’s organic phases, followed by vortex and centrifugation. The organic phase was transferred to a new tube and the remaining half of the samples were spiked ($n = 3$; referred to as the “post-spike samples”). All samples were completely dried under a light stream of nitrogen gas at 37 $^{\circ}$ C. Silylation was performed with 50 μ L BSTFA with 1% TMCS and 50 μ L pyridine, with incubation at 60 $^{\circ}$ C for 60 min. Hereafter, samples were allowed to cool to room temperature and transferred to a GC sample vial with an insert and again capped.

Method 2 (extraction solvent evaluation)²²

This method was identical to that of Method 1, except for the addition of 1.5 mL ice-cold acetonitrile with ethylacetate: Six drops of 5N HCl were added to 100 μ L urine filled up to 500 μ L with a 9% saline solution ($n = 6$). Hereafter, the IS (100 μ L, final concentration of 50 ppm) was added and half of the samples ($n = 3$) were spiked (“pre-spike samples”). Ethylacetate (6 mL) was added to all samples, followed by ice-cold acetonitrile (1.5 mL), whereafter samples were mixed (30 min) and centrifuged (845xg for 5 min at 4 $^{\circ}$ C). After the organic phase was collected, diethylether (3 mL) was added to the remaining aqueous phase, and the samples were again mixed,

centrifuged, and the organic phase collected and combined with the previously collected phase. Approximately 3 g Na_2SO_4 was added to the organic phases of each sample, followed by vortex and centrifugation. The organic phases were then transferred to a new tube and the remaining half of the samples were spiked ($n = 3$; “post-spike samples”). All samples were evaporated to complete dryness under a light stream of nitrogen gas at 37 °C. Silylation was performed with BSTFA with 1% TMCS (50 μL) and pyridine (50 μL), with incubation at 60 °C for 60 min. Hereafter, samples were allowed to cool to room temperature and transferred to a GC sample vial with an insert and again capped.

Method 3 (one-step versus two-step derivatization)²³

This method was identical to that of Method 1, except for an addition methoximation step prior to silylation (a two-step derivatization): After the addition of six drops of 5N HCl and a 50 ppm IS (100 μL) to urine samples ($n = 6$; 100 μL filled up to 500 μL with a 9% saline solution), half of the samples ($n = 3$) were spiked (“pre-spike samples”). Hereafter, and ethylacetate (6 mL) was added to all samples, followed by mixing (30 min), centrifugation (845xg for 5 min at 4 °C), and organic phase collection. Diethylether (3 mL) was added to the remaining aqueous phase, and the samples were again mixed, centrifuged, and the second organic phase collected and combined with the first. Na_2SO_4 (approximately 3 g) was added to the collected organic phases, followed by vortex and centrifugation. Hereafter, the samples were transferred to a new tube and the remaining half were spiked ($n = 3$; “post-spike samples”). All samples were completely dried under a light stream of nitrogen gas at 37 °C prior to methoximation. Thus, 50 μL MOX-HCl dissolved in pyridine (20 mg/mL) was added to each sample, capped, and incubated at 50 °C for 90 min. Hereafter, the samples were left to cool to room temperature prior to silylation with 50 μL BSTFA with 1% TMCS (no additional pyridine was added), with incubation at 60 °C for 60 min. The samples were again allowed to cool to room temperature and transferred to a GC sample vial with an insert and re-capped.

Method 4 (combination of Methods 2 and 3)

This method was identical to that of Methods 2 and 3, combining the use of ice-cold acetonitrile with ethylacetate, and two-step derivatization: To each urine sample ($n = 6$; 100 μL urine filled up to 500 μL with a 9% saline solution), six drops of 5N HCl and 100 μL of an IS (500 ppm) was added. Hereafter, half of the samples were spiked ($n = 3$; “pre-spike samples”). Ethylacetate (6 mL) and ice-cold acetonitrile (1.5 mL) was added to all sample, followed by mixing (30 min), centrifugation (845xg for 5 min at 4 °C), and organic phase collection. Diethylether (3 mL) was then added to the remaining aqueous phase, again followed by mixing, centrifugation, and the organic phase collection (which was combined with the previously collected phase). Na_2SO_4 (approximately 3 g) was added to the collected organic phases, and after vortex and centrifugation, the samples were transferred to a new tube. The remaining half of the samples were then spiked ($n = 3$; “post-spike samples”). All samples were dried with nitrogen gas at 37 °C. Once samples were completely dry, methoximation was performed using 50 μL MOX-HCl dissolved in pyridine (20 mg/mL). Samples were capped and incubated at 50 °C for 90 min. Hereafter, the samples were left to cool to room temperature prior to silylation with 50 μL BSTFA with 1% TMCS (no additional pyridine was added), with incubation at 60 °C for 60 min. The samples were again allowed to cool to room temperature, before transferring them to a GC sample vial with an insert, and re-capped.

Method 5 (DA)¹³

The IS (100 μL ; final concentration of 50 ppm) was added to 100 μL urine. Half of the samples were spiked ($n = 3$; “pre-spike samples”), ice-cold acetonitrile (300 μL) was added, and samples were centrifuged at room temperature (15000xg for 5 min). The supernatant was collected and transferred to a GC–MS vial, whereafter the remaining half of the samples were spiked ($n = 3$; “post-spike samples”). Similar as before, the samples were dried completely under a light stream of nitrogen gas at 37 °C. The vials were allowed to cool to room temperature before applying two-step derivatization: Samples were methoxylated using 50 μL MOX-HCl dissolved in pyridine (20 mg/mL), capped, and incubated at 50 °C for 90 min. After cooling to room temperature, samples were silylated with 50 μL BSTFA with 1% TMCS (no additional pyridine was added), capped, and incubated at 60 °C for 60 min. Samples were transferred to a glass insert, placed into the original GC–MS vial, and capped.

Method 6 (additional drying step)¹⁵

For this part of the investigation, 18 QC aliquots were used, divided into two groups: Group 1 samples ($n = 9$) were subjected to Method 5 exactly as indicated above (including two-step derivatization). Group 2 samples ($n = 9$) were also subjected to Method 5, however, the derivatization process was adjusted by adding another drying step prior to silylation. Following methoximation as described above (50 μL MOX-HCl dissolved in pyridine [20 mg/mL]), samples were again dried under a light stream of nitrogen gas at 37 °C prior to silylation with 50 μL BSTFA with 1% TMCS (no additional pyridine was added). After cooling to room temperature, samples were transferred to a glass insert, placed into the original GC–MS vial, and capped.

Method 7 (urease pre-treatment)

QC aliquots ($n = 30$) were divided into five groups, namely Group 1 urease-treated samples subjected to heating ($n = 6$; UT-H); Group 2 urease-non-treated samples subjected to heating ($n = 6$; NT-H); Group 3 water-treated samples subjected to heating to serve as a blank sample for investigating the diluting effect urease pre-treatment has on samples ($n = 6$; WT-H); Group 4 urease-treated samples subjected to sonication instead of heating ($n = 6$; UT-S); Group 5 urease-non-treated samples subjected to sonication ($n = 6$; NT-S). The analysis of these samples was based on Method 6, however, following the addition of the IS, urease (100 μL of a 1 mg/mL solution; Groups 1 and 4) or water (100 μL ; Group 3) was also added. The samples were then either heated (30 min at

37 °C; Groups 1–3)¹⁶ or sonicated (30 min at room temperature with 005 power; Groups 4 and 5)¹⁸ for adequate reaction to occur, before continuing with the addition of ice-cold acetonitrile whereafter the rest of the method proceeded as described previously (see “Method 6”).

GC–MS analysis

Prior to sample analysis, a routine clean-up and maintenance check was performed (i.e., leak check, tune check, and mass calibration) and a new liner and septum was inserted to prevent any undesired reactions and surface adsorption phenomena.

Samples were analysed in two batches (Part A; Methods 1–5) and (Part B; Methods 6 and 7), in a randomised manner. The samples were placed into an auto-sampler tray (Gerstel MS Germany), combined with a Pegasus 4D GCxGC-TOFMS system (LECO Africa [Pty] Ltd, Johannesburg, South Africa) and fitted with an Agilent 7890 GC and TOFMS (LECO Africa). All samples (1 µl) were injected using a 1:10 split ratio with purified helium as a carrier gas, set at a constant flow of 1.4 mL/min. Chromatographic separation was achieved with a Restek Rxi-5MS primary capillary column (28.2 m; 250 µm diameter; 0.25 µm film thickness), and a Restek Rxi-17 secondary capillary column (1.3 m; 250 µm diameter; 0.25 µm film thickness). Throughout the entire chromatographic run, the front inlet temperature was maintained at 250 °C, the transfer line at 225 °C, and the ion source was kept constant at 200 °C. Cryomodulation was achieved with a hot pulse of nitrogen gas for 0.7 s, every 3 s. The primary oven programme (see Figure S1 in the supplementary information) started at 70 °C, which was held for 1 min. The temperature was then ramped up as follows: 5 °C/min to 100 °C, 10 °C/min to 160 °C, 13 °C/min to 230 °C, and finally 20 °C/min to 300 °C, which was held for 2 min. The secondary oven was programmed identical to that of the primary oven, except for a + 5 °C at each interval. The total run time of each sample was ≈24 min. Prior to MS data acquisition, a 480 s solvent delay was used, during which no mass spectra were captured. However, to accurately represent retention times, this interval was appended to the time axis of the GC column. Mass spectra were collected over a range of 50–800 m/z at an acquisition rate of 20 spectra per second. The filament bias was -70 eV and the detector voltage was optimised at an offset of 50 V.

Data analysis

For each experimental dataset, Leco® ChromaTOF-GC Software (v4.72.0.0) was used for MS deconvolution (with a signal-to-noise ratio of 100 and a minimum of three apexing peaks), peak identification (with a 60% similarity match) and peak alignment. The raw data from the GCxGC-TOFMS was exported as a .csv file and converted to an Excel Workbook (.xlsx) format. The data was normalised based on the IS, followed by normalisation based on the creatinine value. Various parameters were assessed using Excel, including analytical repeatability, metabolome coverage, and metabolite recovery.

Results and discussion

In Part A of this investigation, five methods were evaluated based on repeatability, metabolome coverage, and metabolite recovery. The goal here was to select the most appropriate method for non-targeted GC–MS application. Hereafter, the selected method was further explored for possible optimisation in Part B.

Repeatability

To determine the analytical repeatability for Methods 1–5, the first three samples of each method (i.e., the “pre-spike samples”) were used to calculate the CV values, expressed as percentage. When interpreting the CV values, the more overall compounds with a CV < 50%, the better the method²⁴. Figure 2a illustrates the CV distribution for all detected compounds, for the five extraction methods. When considering the analytical repeatability, Method 5 had the best repeatability (82.4% of all detected compounds had a CV < 50%), followed by Method 3 (62.1%), Method 4 (61.9%), Method 1 (54.6%), and lastly Method 2 (36.4%).

For further confirmation of the above ranking, the web-based server *MetaboAnalyst* was used to statistically evaluate the data and compile a principal component analysis (PCA) scores plot (see Fig. 2b). This plot provides an overview of the overall similarity or dissimilarity among all five extraction methods without directly incorporating the CV values, but rather considering patterns and relationships among the samples of each method based on their measured variables. Indeed, the PCA analysis revealed distinct clustering patterns among the five methods. Specifically, the OA extraction methods (Methods 1–4) exhibited a tight clustering, suggesting similar metabolite profiles. The DA samples (Method 5) were positioned distinctly and significantly distant from the OA extraction methods. This observation suggests substantial dissimilarity in the metabolite profiles of Method 5 compared to the other methods, likely due to its unique extraction principle and variability in its metabolite composition. On a PCA, the approach with the least amount of variance will have the smallest 95% confidence region circle, which is where the samples cluster closest together. Thus, the bigger the 95% confidence region, the less repeatable the method. The large green ellipse, which represents Method 2 (i.e., the addition of acetonitrile), shows great variance between the samples. Although acetonitrile has a superior extraction efficiency, Majors (2013)²² also observed that its extracts were still somewhat “dirty”, possibly contributing to the observed variation. Furthermore, the presence of water in acetonitrile is a potential factor in the variance between the extracted analytes, as it may impact how effectively different OAs are extracted, depending on how the OA is distributed between the aqueous and organic phases²⁵. This may lead to the loss of compounds during extraction, making this method the least repeatable. Thus, it may be beneficial to add kosmotropic salts (e.g., MgSO₄ or NaCl) to phase separate acetonitrile from a water/acetonitrile solution. The original in-house method samples (Method 1; red ellipse), has slight variation between the samples but less so compared to Method 2. Considering the additional compound stabilisation that oximation offers²³, Method 3 (dark blue ellipse) and Method 4 (blue ellipse) have better confidence region circles than Methods 1 and 2, indicating less variability, which concurs with Fig. 2a.

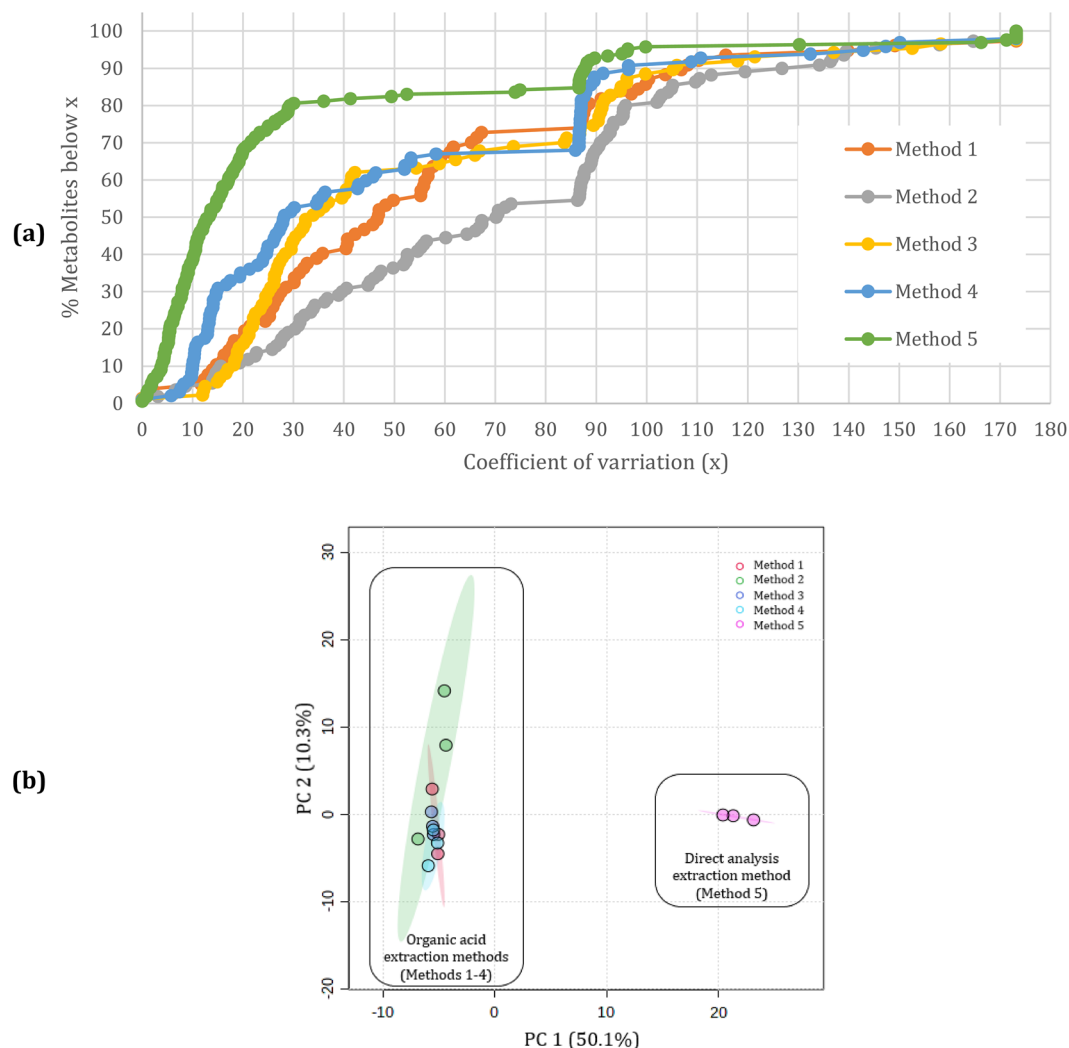


Figure 2. Evaluation of repeatability of the five extraction methods. **(a)** The percentage of metabolites below a coefficient of variation (CV) threshold < 50%, ranked as Method 5 (82.4%), Method 3 (62.1%), Method 4 (61.9%), Method 1 (54.6%), and Method 2 (36.4%). **(b)** Principal component analysis (PCA) scores plot which illustrates the clustering patterns among the samples based on their metabolite profiles/composition. The plot provides an overview of the overall similarity or dissimilarity among the methods' samples without directly incorporating the CV values.

Similarly, Method 5 (pink ellipse), which also entails oximation prior to silylation, shows tight sample clustering. Thus, the method ranking based on the PCA scores plot corresponds with that of the CV-graph.

Based on Fig. 2, Method 5 yields more repeatable results compared to any of the OA extraction methods. This is possibly due to oximation for compound stabilisation, in conjunction with its easy preparation (i.e., reduced analytical steps involved) since the OA extraction methods are more labour-intensive and hence more prone to errors.

Metabolome coverage

The metabolome coverage of each method was assessed using the normalised “pre-spike samples” only. The total number of compounds detected across all five extraction methods ($n = 220$) was determined, whereafter the number of overall compounds not detected in at least 50% of all samples for each extraction method was calculated as: % compounds undetected = $\frac{\text{Number of compounds with zero average}}{220 \text{ compounds overall}} \times 100$. The extraction methods were then ranked based on this calculated percentage, from the least number of undetected compounds to the most undetected compounds as follows: Method 5 (26.4% undetected); Method 2 (51.4% undetected); Method 4 (57.3% undetected); Method 3 (61.8% undetected); and Method 1 (66.4% undetected). Upon further investigation, we found that 91 metabolites were exclusively detected using Method 5, whereas 58 were specific to the four OA extraction methods. Of these 58, 13 were exclusive to Method 2, seven were exclusive to Method 4, one was exclusive to Method 3, and none was exclusive to Method 1 (see Table 1 for complete lists of these unique

Method name and number of unique compounds	List of unique compounds			
Method 1 (n = 0)	None			
Method 2 (n = 13)	Analyte 003 Analyte 013 Analyte 121 Analyte 128	Analyte 135 Analyte 148 Analyte 150	Analyte 156 Analyte 157 Analyte 167	Analyte 171 Analyte 192 Analyte 308
Method 3 (n = 1)	Analyte 322			
Method 4 (n = 7)	2,3,4-Trihydroxy-butylaldehyde O-methyl-oxime	Analyte 036 Analyte 071	Analyte 209 Analyte 216	Analyte 325 Malonic acid
Method 5 (n = 91)	2-Amino-benzoic acid 3,4,5-Trihydroxy-pentanoic acid 3,4-Dihydroxy-butyric acid 3R,4S-Tetrahydrofuran-2,3,4-triol 4-Hydroxy-pyrrolidine-2-carboxylic acid 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [1] 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [2] 5-Oxo-pyrrolidine-2-carboxylic acid Analyte 004 Analyte 005 Analyte 023 Analyte 025 Analyte 029 Analyte 035 Analyte 038 Analyte 043 Analyte 046	Analyte 050 Analyte 057 Analyte 058 Analyte 061 Analyte 063 Analyte 064 Analyte 070 Analyte 072 Analyte 085 Analyte 086 Analyte 087 Analyte 090 Analyte 096 Analyte 099 Analyte 118 Analyte 125 Analyte 126 Analyte 131 Analyte 137 Analyte 145 Analyte 151 Analyte 177 Analyte 183 Analyte 186 Analyte 199	Analyte 210 Analyte 212 Analyte 218 Analyte 220 Analyte 225 Analyte 251 Analyte 253 Analyte 266 Analyte 276 Analyte 278 Analyte 280 Analyte 285 Analyte 287 Analyte 302 Analyte 305 Analyte 306 Analyte 311 Analyte 321 Analyte 323 Arabinofuranose Arabinose Ascorbic acid Aucubin Butane-2,3-diol Creatinine	Dehydroabietic acid Erythronic acid γ -lactone Galactose [1] Galactose [2] Glucose Glucuronic acid Glucuronic acid γ -lactone Glycerol Isocitric lactone meso-Erythritol Methyl galactoside Methylmalonic acid Myo-Inositol [1] Myo-Inositol [2] N-Acetyltyrosine N-benzoyl-glycine Oxalic acid Oxalic acid [1] Oxalic acid [2] Rhamnose Ribofuranose Tagatofuranose Threo-3-deoxy-pentonic acid Xylose
Compounds unique to the organic acid methods (Methods 1–4)* (n = 58)	2,3,4-Trihydroxy-butylaldehyde O-methyl-oxime 2,4-Dihydroxybutyric acid 2-Hydroxy-pentanedioic acid 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [1] 5-Hydroxymethyl-tetrahydro-furan-2,3,4-triol [2] Analyte 003 Analyte 013 Analyte 014 Analyte 020 Analyte 022 Analyte 024	Analyte 027 Analyte 031 Analyte 036 Analyte 049 Analyte 052 Analyte 055 Analyte 059 Analyte 071 Analyte 121 Analyte 123 Analyte 128 Analyte 135 Analyte 143 Analyte 144 Analyte 148 Analyte 150	Analyte 156 Analyte 157 Analyte 167 Analyte 168 Analyte 170 Analyte 171 Analyte 173 Analyte 174 Analyte 175 Analyte 176 Analyte 192 Analyte 195 Analyte 205 Analyte 209 Analyte 216 Analyte 226	Analyte 239 Analyte 240 Analyte 255 Analyte 297 Analyte 301 Analyte 308 Analyte 314 Analyte 322 Analyte 325 Benzene-1,2,3-triol Citric acid [miss ID] Malonic acid diethyl ester Octadecanoic acid Propane-1,2,3-triol [1] Tetrahydro-pyran-2,3,4,5-tetraol (isomer 2)

Table 1. Unique compounds detected for each method, and those detected in only the organic acid extraction methods. *Some of these compounds were detected across multiple organic acid extraction methods, hence they are not unique to a specific organic acid method (Methods 1–4), however, they were not detected using the direct analysis method (Method 5).

compounds). As can be seen, both metabolome coverage rankings (either considering the percentage of undetected compounds or the number of unique compounds) concurred.

Metabolite recovery

A spiking mixture is a carefully prepared solution containing known concentrations of specific compounds (see Section "Standards"), and serves as a vital component in analytical chemistry, fulfilling various purposes. In this study, a spiking mixture was used to specifically assess the recovery rate of target analytes during sample preparation and analysis. This assessment involves adding a known amount of analyte into the natural test sample matrix. By comparing the amount of spike compounds added to the sample with the amount detected in the final analysis, researchers can calculate the recovery percentage, which indicates the efficiency of target analyte extraction and measurement. Thus, metabolite recovery was determined by comparing the concentrations of compounds in the "pre-spike" and "post-spike" samples, using only those compounds included in the spiking solution (see Section "Standards"). The percentage of recovery was calculated as: % recovery = $\frac{\text{Pre-spike concentration}}{\text{Post-spike concentration}} \times 100$ (see Table 2).

Among the evaluated methods, Method 5 exhibited an average percentage recovery of 57.7%. Notably, citric acid displayed the highest recovery (118.1%), followed by phenylalanine (86.6%), ascorbic acid (86.5%), succinic acid (78.8%), palmitic acid (69.3%), alanine (61.4%), leucine (45.7%), and GABA (30.5%). For Method 2, the average recovery percentage was 35%, with the highest recovery observed for palmitic acid (181%), followed by

Spiking compound	Method 1 (original in-house)			Method 2 (acetonitrile addition)			Method 3 (two-step derivatization)			Method 4 (combined methods 2 & 3)			Method 5 (direct analysis)		
	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery	Pre-spike average	Post-spike average	% Recovery
4-Aminobutyric acid (GABA)	0.00	0.74	0.00	0.00	0.95	0.00	0.00	0.84	0.00	0.00	0.23	0.00	0.11	0.35	30.46
L-Alanine	0.00	3.06	0.00	0.00	4.01	0.00	0.00	3.70	0.00	0.00	1.79	0.00	1.85	3.01	61.35
L-Leucine	0.00	1.54	0.00	0.00	0.18	0.00	0.00	1.47	0.00	0.00	0.23	0.00	0.45	0.99	45.66
L-Tryptophan	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
L-Tyrosine	0.00	0.28	0.00	0.00	0.24	0.00	0.00	0.37	0.00	0.00	0.17	0.00	0.00	0.03	0.00
L-Phenylalanine	0.00	0.53	0.00	0.00	0.65	0.39	0.09	0.78	11.94	0.03	0.45	6.15	0.21	0.24	86.55
Ascorbic acid	0.00	0.02	0.00	0.00	0.04	0.00	0.00	0.09	0.00	0.00	0.01	0.00	0.09	0.11	86.49
Citric acid	0.57	0.79	72.25	1.02	1.24	82.56	0.59	0.89	65.61	0.90	0.69	130.81	1.63	1.38	118.12
Succinic acid	0.11	0.13	82.33	0.13	0.15	85.62	0.11	0.14	80.08	0.10	0.12	79.05	0.16	0.21	78.77
Palmitic acid	0.14	0.21	67.57	0.38	0.21	181.00	0.17	0.34	50.67	0.12	0.18	67.16	0.09	0.13	69.28
Average recovery			22.22			34.96			20.83			28.32			57.67

Table 2. Metabolite recovery of the compounds in the spiking mixture (containing pure, known standard concentrations of different class compounds) indicated for each method. The compounds are grouped per class.

succinic acid (85.6%), citric acid (82.6%), and phenylalanine (0.4%). Method 4 achieved an average recovery percentage of 28.3%, with citric acid showing the highest recovery (130.8%), followed by succinic acid (79.1%), palmitic acid (67.2%), and phenylalanine (6.2%). In the case of Method 1, the average percentage recovery was 22.2%, with succinic acid demonstrating the highest recovery (82.3%), followed by citric acid (72.3%) and palmitic acid (67.6%). Method 3 exhibited the lowest average percentage recovery at 20.8%, with succinic acid achieving the best recovery (80.1%), followed by citric acid (65.6%), palmitic acid (50.7%), and phenylalanine (27%). It is noteworthy that some compounds showed recoveries above 100%. This can be attributed to the fact that the QC samples used during analysis were compiled from human urine rather than synthetic urine, ensuring a realistic representation in terms of robustness during this investigation. Thus, many of these compounds may already be present in the samples.

Alanine, GABA, ascorbic acid, and leucine exhibited poor extraction efficacy across all methods, resulting in a 0% recovery, except for Method 5 (DA), which extracted all these metabolites. It is well-known that the polarity of OAs can significantly influence their extraction efficiency. Acidic compounds containing more hydroxyl groups tend to have lower recovery efficacy, as observed with ascorbic acid, which is very acidic with four hydroxyl groups. Mouskeftara et al. (2021)²⁶ also reported significantly lower GABA signals when using ethyl acetate for analyte extraction compared to an acidic extraction with methyl tert-butyl ether (MTBE) or a more commonly used approach involving methanol (MeOH). Their study suggested that a combination of MeOH and methoxyamine with N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) 1% TMCS derivatization generally enhanced efficiency. However, BSTFA 1% TMCS exhibited a preference for certain acids, while MSTFA 1% TMCS yielded lower peaks for succinic acid. In contrast, our study achieved a recovery efficiency of 79% or more across all the OA extraction methods using BSTFA 1% TMCS. For future studies employing an LLE approach, exploring the potential use of MTBE in combination with acetonitrile is worth considering. Additionally, Pasikanti et al. (2008)²⁷ compared various derivatization agents [BSTFA, MSTFA, and N-methyl-bis(trifluoroacetamide) (MBTFA)] used in urinary GC–MS metabolomics. In terms of the number of observed peaks, peak intensity, and reproducibility, both BSTFA and MSTFA demonstrated equal derivatization efficiency. Since MBTFA is more effective in derivatizing secondary and tertiary amine groups, they also explored the combination use of MBTFA with BSTFA. When MBTFA derivatization was preceded by BSTFA derivatization, more amino acids were detected, however, the integrated peak areas had poor reproducibility. Based on these findings, the authors recommended a preference for BSTFA as the derivatization agent.

Tryptophan and tyrosine had 0% recovery across all extraction methods, however, this was expected since GC systems are not ideal for detecting amino acids with great precision^{25,28}. By comparing the various OA extraction techniques, Methods 3 and 4 were able to recover the most amounts of phenylalanine. This is most likely due to the addition of an oximation step, which produces methoxime derivatives that are more stable and improves peak identification^{26,29,30}. The metabolite recovery ranking, however, shows that acetonitrile is required for higher extraction efficacy²² and that oximation alone is insufficient to recover broad ranges of metabolite classes.

However, it is important to exercise caution when interpreting metabolomics results, with particular attention to the specific methodology employed for sample extraction, as well as the analytical platform used for

	Repeatability	Metabolome coverage	Metabolite recovery	Cumulative score
Method 1 (original in-house)	4	5	4	13
Method 2 (acetonitrile addition)	5	2	2	9
Method 3 (two-step derivatization)	2	4	5	11
Method 4 (combination of Methods 2 & 3)	3	3	3	9
Method 5 (direct analysis)	1	1	1	3

Table 3. Method rankings based on validation parameters, and an overall cumulative score of each.

analysis. In the field of metabolomics, the choice of analytical platform is a critical consideration when designing a study. Each platform possesses its unique strengths and limitations in terms of detecting specific metabolites or metabolite classes. For instance, nuclear magnetic resonance (NMR) is particularly adept at detecting sugars, while liquid chromatography (LC) platforms excel in the analysis of amino acids, which gas chromatography (GC) systems may have limitations in detecting with high precision. However, GC platforms offer high efficiency, sensitivity, and reproducibility, making them well-suited for conducting non-targeted or so-called "discovery mode" metabolomics research where a broad range of metabolites needs to be analysed. In the context of this study, our focus was primarily on the application of methods optimised for a GC–MS platform, which has its own unique advantages. Researchers should carefully select the analytical platform that aligns with the specific metabolites of interest and the objectives of their metabolomics study to ensure meaningful and accurate results. While our methods may not be universally applicable, they provide valuable insights into the efficient extraction and analysis of metabolites using a GC–MS platform, and future studies exploring other platforms for specific metabolites or metabolite classes are certainly warranted.

Selecting the most appropriate method for non-targeted GC–MS application

After each validation parameter was investigated, each method was ranked based on its overall performance, and a cumulative score was calculated for each method (see Table 3). Since a top ranking would result in a lower cumulative score, the method with the lowest cumulative score should theoretically be the superior method.

The ranking provided here can be considered for future studies related to semi-targeted metabolomics³¹. It should be noted that repeatability is subjective to the analyst's ability. Naturally, the fewer analytical steps involved, the simpler the method would be, and hence perform better in terms of repeatability. When an investigation requires a semi-targeted approach to extract OAs, for example, Methods 2 or 4 should be considered for optimal results. Furthermore, both options allowed for a high number of unique compounds and the overall recovery of organic and fatty acids was good for both methods. On the other hand, a non-targeted strategy examines the complete metabolic state³² and would therefore potentially benefit more from taking the percentage of undetected compounds into account, as the investigation would need to identify multiple groups of substances or different compound classes simultaneously³³.

Considering the aim of the investigation, i.e., to optimise a low-volume urine preparation procedure for non-targeted GC–MS analyses covering various compound classes, Method 5 outperformed all other methods, consistently ranking first in every validation parameter. Therefore, we attempted to further optimise Method 5 by determining if an additional drying step between two-step derivatization (Method 6), and urease sample pre-treatment, also considering the effect of heating versus sonication (Method 7), is beneficial.

Additional drying between two-step derivatization

For this part of the investigation, samples were extracted as described in Section "Method 6 (additional drying step)". When considering the total ion chromatograms (Fig. 3) from Sample A (orange; traditional two-step derivatization) and Sample B (green; additional drying between two-step derivatization), the additional drying step seems to reduce the signal of urea significantly, which is considered advantageous.

Although the average peak height was visually higher in the conventional two-step derivatization samples, certain metabolite classes still displayed higher reported areas. One such group was amino acids, which has also been reported by Liebeke and Puskás (2019)¹⁵ to be higher following an additional drying step samples. Moreover, the inclusion of an additional drying step results in a reduction of signal for specific metabolites often referred to as "junk compounds", which commonly arise from reagent artifacts.

Urease sample pre-treatment

To evaluate the potential advantages of urease sample pre-treatment and examine the impact of heating versus sonication, sample analysis was conducted following the procedure outlined in Section "Method 7 (urease pre-treatment)". The five previously mentioned sample groups (UT-H, NT-H, WT-H, UT-S, and NT-S) were compared in two separate experimental sets. The previously described criteria of repeatability, metabolome coverage, and metabolite recovery (Sections "Repeatability", "Metabolome coverage" and "Metabolite recovery") were employed to assess the outcomes of both experiments. Firstly, the effect of urease pre-treatment (during which samples were subjected to heating only, i.e., UT-H, NT-H, and WT-H) was evaluated. The results obtained in every validation parameter is indicated in Table 4. Here, the water-treated samples had the lowest cumulative ranking score between the three groups, and hence performed the best, with the urease-treated samples having the most unfavourable ranking in each assessment.

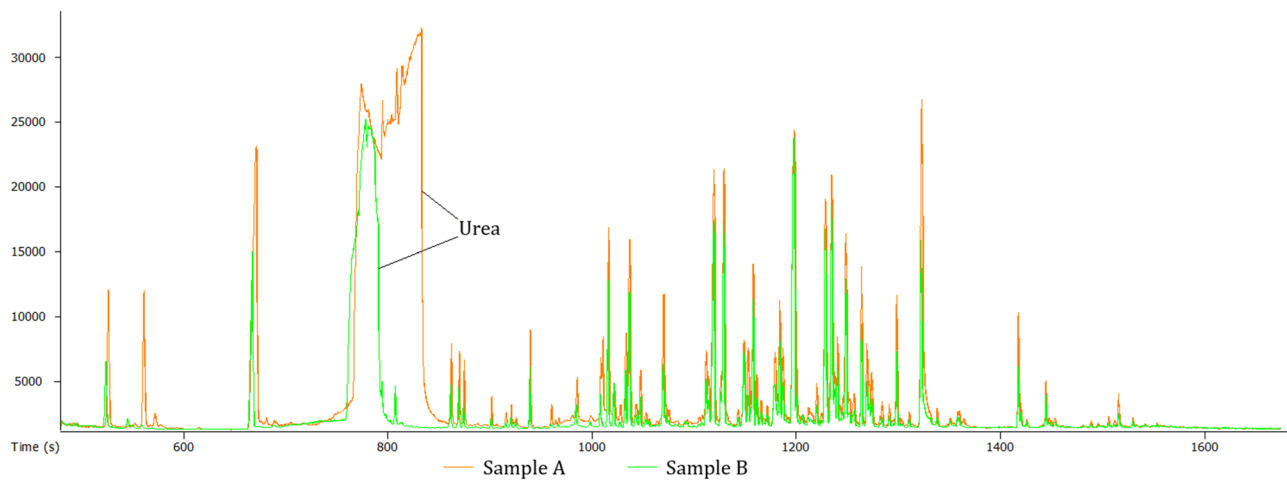


Figure 3. The total ion chromatograms (TICs) showing the effect of an additional drying step. Sample A (orange) was subjected to two-step derivatization while Sample B (green) was subjected to an additional drying step between two-step derivatization. Sample B showed a significant reduction in the signal of urea, which is considered advantageous.

	Repeatability	Metabolome coverage	Metabolite recovery	Cumulative score
Urease-treated samples (UT-H)	3	3	3	9
Urease-non-treated samples (NT-H)	2	2	2	6
Water-treated samples (WT-H)	1	1	1	3

Table 4. Methods rankings based on validation parameters, when comparing urease-treated, urease-non-treated, and water-treated samples.

Urease treatment effectively reduces the urea content in urine samples, but it can have additional effects beyond urea reduction, which can negatively impact sample performance. These effects are influenced by various factors. Firstly, undesired enzymatic reactions may occur since enzymes can catalyse reactions beyond the intended target, leading to chemical changes in the sample. These changes may include alterations in the metabolite profiles, degradation of certain compounds, or the generation of new metabolites as by-products of the enzymatic reactions. Such effects can impact the composition, stability, and integrity of the metabolites of interest, introducing variability and decreasing the accuracy of the analysis²⁰.

Secondly, urease treatment may decrease the recovery of metabolites of interest due to chemical transformations or interactions with the enzymatic products generated during urea hydrolysis. If certain metabolites are poorly recovered or lost entirely due to urease treatment, it can result in gaps or biases in the metabolome coverage, compromising the accuracy and reliability of the analysis. Consequently, a comprehensive understanding of the metabolic profile may be hindered, potentially leading to incomplete or misleading interpretations of the results^{7,16}.

Thirdly, water-treated samples may exhibit a slight dilution effect, enhancing the extraction capability and improving the detection and measurement of metabolites¹⁶. Dilution helps mitigate potential matrix effects or sample interferences that could affect the analysis. Matrix effects can occur when the components of the sample, such as salts or proteins, interfere with the detection and measurement of analytes. By diluting the sample, the concentrations of these interfering components are reduced, potentially minimizing their impact on the analysis³⁴. Consequently, a more accurate and reliable measurement of the target analytes can be achieved. However, the optimal concentration range depends on the specific technique and analyte characteristics, as extremely high or low concentrations may present challenges in terms of sensitivity, linearity, or dynamic range.

Fourthly, urease activity itself can cause interference or unwanted changes in the sample due to its inherent catalytic activity^{35,36}. This interference can manifest as chemical interactions or alterations in the stability, integrity, or concentration of metabolites of interest.

Lastly, urease treatment adds complexity and additional steps to the sample preparation process, increasing the risk of errors or variability. Each additional step presents an opportunity for introducing artifacts, contamination, or analyte loss¹⁶. Inaccurate or inconsistent handling, mixing, or transferring of samples during urease pre-treatment can impact the accuracy, precision, and reproducibility of the analysis.

The total ion chromatograms (Fig. 4) of a urease-treated sample (Sample A; green) and a urease-non-treated sample (Sample B; orange) were compared. Based on visual inspection, urease pre-treatment indeed reduced the amount of urea significantly. However, urease also reduced the overall metabolic signals, likely due to the various factors discussed above. This again highlights the detrimental effect that urease pre-treatment can have

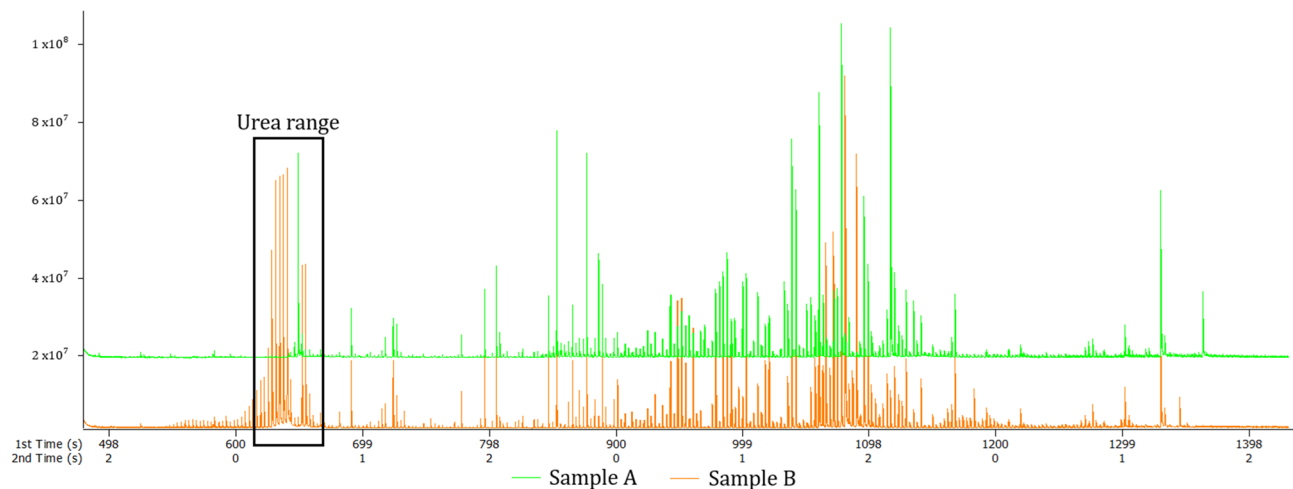


Figure 4. The total ion chromatograms showing the effect of urease sample pre-treatment. Sample A (green) was subjected to urease pre-treatment while Sample B (orange) shows a urease-non-treated sample.

	Repeatability	Metabolome coverage	Metabolite recovery	Cumulative score
Urease-treated: Heating (UT-H)	4	4	4	12
Urease-treated: Sonicate (UT-S)	2	3	2	7
Urease-non-treated: Heating (NT-H)	1	2	3	6
Urease-non-treated: Sonicate (NT-S)	3	1	1	5

Table 5. Method rankings based on validation parameters, when comparing heating and sonication following urease pre-treatment.

when non-targeted metabolomics studies are performed. In disease characterisation studies, for example, the metabolic profiles of patients need to remain unchanged, hence urease pre-treatment would not be recommended for such patient samples.

Hereafter, an additional experiment was conducted to examine the impact of heating versus sonication during urease pre-treatment to determine if this could increase the benefits of such an approach. Thus, UT-H, NT-H, UT-S, and NT-S were compared (Section "Method 7 (urease pre-treatment)") and again assessed based on the previously described criteria of repeatability, metabolome coverage, and metabolite recovery (Sections "Repeatability", "Metabolome coverage" and "Metabolite recovery"), results of which are indicated in Table 5.

Based on the results, an approach of urease non-treatment, regardless of heating or sonication, consistently outperforms approaches where samples are subjected to urease pre-treatment, demonstrating superior performance, and yielding more reliable results. These findings align with the conclusions reached by Kim et al. (2020)²⁰, who observed that urease pre-treatment may not be as advantageous as initially believed. Their study revealed that urease pre-treatment introduced artefacts into the metabolite profiles, potentially leading to misinterpretation of results. Other studies, however, observed that urease pre-treatment samples exhibited higher concentrations compared to urease non-treated¹⁸ and water-treated samples¹⁶. Nonetheless, Palmas et al. (2018)¹⁸ observed a small number of metabolites in urease non-treated samples that exhibited higher concentrations compared to the sonicated urease-treated samples, aligning with our findings. Although the exact origin remains unknown, they postulated that some of the ammonia generated during the urea conversion process in sonication may influence the concentrations of various metabolites. When urea undergoes conversion into ammonia and carbon dioxide, the former may react with acidic species, resulting in the formation of ammonium salts. Since GC-MS cannot detect salts, it may explain the reduced metabolite levels. Furthermore, this study noted that a longer sonication duration (1-h) combined with 200 µL urease solution is optimal, suggesting a potential avenue for future urease pre-treatment studies involving sonication.

Considering the continued contradictory results regarding urease pre-treatment, researchers should carefully assess its effects and evaluate its potential benefits primarily for specific analyses focused on urea-related measurements, while also ensuring the reliability and validity of the results. It is important to evaluate the trade-off between urea reduction and the potential compromise in metabolite recovery to ensure the most accurate and comprehensive analysis of the metabolome. If urease pre-treatment is deemed necessary, sonication offers numerous advantages in terms of efficiency, homogeneity, accessibility, avoidance of thermal degradation, and versatility, making it a preferred option to heating.

In terms of heating versus sonication when performing urease pre-treatment, results indicate that sonication is the preferred method to catalyse this reaction. Sonication can be more advantageous than heating due to several reasons, including (i) efficient and rapid mixing and agitation of the sample; (ii) achieving a more homogeneous distribution of the urease enzyme throughout the sample; (iii) facilitating the penetration of

the enzyme into cellular compartments or structures, which may be more difficult to access via heating alone; and (iv) avoidance of thermal degradation of sensitive compounds present in the sample as it does not involve significant temperature increases^{37,38}.

Implications of the findings

The implications of our findings are significant for clinical, nutritional, and biological investigations focused on examining metabolic abnormalities associated with various pathophysiological conditions. Depending on the specific research objectives, our extraction methods can serve as valuable guidelines. Disease characterisation, for example, typically employs non-targeted metabolomics approaches (such as Method 6) to obtain comprehensive metabolic information about individuals³⁹. The data analysis can lead to the development of new hypotheses, facilitating the exploration of previously undiscovered disease mechanisms, host–pathogen interactions and adaptations²¹. When disease characterisation indicates abnormalities in OA or urea profiles, researchers may opt for a more semi-targeted approach (such as Methods 2 or 4, as per our results) or targeted approach (Method 7 with sonication), respectively.

While this study primarily adopts a “discovery mode” process with non-targeted methods, aimed at comparative analysis between different techniques to identify the superior method, it is important to acknowledge the limitations of a non-targeted approach. A targeted metabolomics approach, while providing more accurate quantitative results, focuses only on a predetermined list of metabolites⁴⁰. The optimised extraction method presented here holds promise for application in clinical laboratories, especially where time and resources are limited. This efficient method minimises extraction solvents, making it suitable for high-throughput analysis. Its potential to enhance treatment strategies, overcome diagnostic limitations, and enable prompt medical intervention in disease cases makes it a valuable tool for researchers seeking to optimise urinalysis and make informed decisions aligned with their research objectives.

Conclusion

Since non-targeted GC–MS metabolomics aims to investigate multiple metabolite classes at various concentrations, optimal compound extraction is essential. Five urine sample preparation methods were evaluated in terms of repeatability, metabolome coverage, and metabolite recovery, with the aim of determining which is better for non-targeted GC–MS metabolomics applications. A low-volume DA method outperformed all other methods, consistently ranking first in every validation parameter. This method identified 91 unique metabolites, from multiple compound classes, and showed high metabolite recovery. An extraction method that can simultaneously identify all the metabolites in a urine sample is not yet available, mostly due to the complexities and the variety of physiochemical properties of metabolites. Nonetheless, this DA method requires little sample volume, have few analytical steps (which would be more time efficient, cost-effective, and repeatable), and can extract as much of the metabolome as possible, making it ideal for non-targeted urine GC–MS analysis.

Furthermore, the DA approach was optimised by incorporating an additional drying step between two-step derivatization but did not benefit from urease sample pre-treatment. Indeed, we recommend that urease pre-treatment should be reserved for specific analyses targeting urea-related measurements. If urease pre-treatment is necessary, sonication is recommended due to its numerous advantages, including efficiency, homogeneity, accessibility, avoidance of thermal degradation, and versatility compared to heating.

Data availability

All data is available on BioStudies (Accession Number [S-BSST1138](#)).

Received: 13 July 2023; Accepted: 11 October 2023

Published online: 16 October 2023

References

- Collino, S., Martin, F. P. J. & Rezzi, S. Clinical metabolomics paves the way towards future healthcare strategies. *Br. J. Clin. Pharmacol.* **75**, 619–629 (2013).
- Isa, F. *et al.* Mass spectrometric identification of urinary biomarkers of pulmonary tuberculosis. *EBioMedicine* **31**, 157–165 (2018).
- Fiehn, O. Metabolomics by gas chromatography–mass spectrometry: Combined targeted and untargeted profiling. *Curr. Protocols Mol. Biol.* **114**, 30.34.31–30.34.32 (2016).
- Christou, C., Gika, H. G., Raikos, N. & Theodoridis, G. GC–MS analysis of organic acids in human urine in clinical settings: a study of derivatization and other analytical parameters. *J. Chromatogr. B* **964**, 195–201 (2014).
- De Villiers, L. & Loots, D. T. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Curr. Metab.* **1**, 306–317 (2013).
- Mamyryn, B. Time-of-flight mass spectrometry (concepts, achievements, and prospects). *Int. J. Mass Spectrom.* **206**, 251–266. [https://doi.org/10.1016/S1387-3806\(00\)00392-4](https://doi.org/10.1016/S1387-3806(00)00392-4) (2001).
- Khodadadi, M. & Pourfarzam, M. A review of strategies for untargeted urinary metabolomic analysis using gas chromatography–mass spectrometry. *Metabolomics* **16**, 1–14 (2020).
- Chan, E. C. Y., Pasikanti, K. K. & Nicholson, J. K. Global urinary metabolic profiling procedures using gas chromatography–mass spectrometry. *Nat. Protoc.* **6**, 1483–1499 (2011).
- Gupta-Wright, A., Peters, J. A., Flach, C. & Lawn, S. D. Detection of lipoarabinomannan (LAM) in urine is an independent predictor of mortality risk in patients receiving treatment for HIV-associated tuberculosis in sub-Saharan Africa: a systematic review and meta-analysis. *BMC Med.* **14**, 53 (2016).
- Madla, S., Miura, D. & Wariishi, H. Optimization of extraction method for GC–MS based metabolomics for filamentous fungi. *J. Microb. Biochem. Technol.* **4**, 5–9 (2012).
- Su, Y., Xia, S., Wang, R. & Xiao, L. Phytohormonal quantification based on biological principles. *Hormone Metab. Signal. Plants* **13**, 431–470 (2017).

12. Soares-da-Silva-Burato, J., Vargas-Medina, D. A., de Toffoli, A. L., Vasconcelos-Soares-Maciel, E. & Mauro-Lanças, F. Recent advances and trends in miniaturized sample preparation techniques. *J. Separat. Sci.* **43**, 202–225. <https://doi.org/10.1002/jssc.201900776> (2020).
13. Esterhuizen, K. *et al.* A urinary biosignature for mitochondrial myopathy, encephalopathy, lactic acidosis and stroke like episodes (MELAS). *Mitochondrion* **45**, 38–45 (2019).
14. Glicksberg, L. & Kerrigan, S. Derivatization. *Principles of forensic toxicology*, 163–175 (2020).
15. Liebeke, M. & Puskás, E. Drying enhances signal intensities for global GC–MS metabolomics. *Metabolites* **9**, 68 (2019).
16. Webb-Robertson, B.-J. *et al.* A statistical analysis of the effects of urease pre-treatment on the measurement of the urinary metabolome by gas chromatography–mass spectrometry. *Metabolomics* **10**, 897–908 (2014).
17. Wells, W., Chin, T. & Weber, B. Quantitative analysis of serum and urine sugars by gas chromatography. *Clin. Chim. Acta* **10**, 352–359 (1964).
18. Palmas, F., Mussap, M. & Fattuoni, C. Urine metabolome analysis by gas chromatography–mass spectrometry (GC–MS): Standardization and optimization of protocols for urea removal and short-term sample storage. *Clin. Chim. Acta* **485**, 236–242. <https://doi.org/10.1016/j.cca.2018.07.006> (2018).
19. Kind, T., Tolstikov, V., Fiehn, O. & Weiss, R. H. A comprehensive urinary metabolomic approach for identifying kidney cancer. *Anal. Biochem.* **363**, 185–195 (2007).
20. Kim, J. *et al.* Systematic re-evaluation of the long-used standard protocol of urease-dependent metabolome sample preparation. *PLoS ONE* **15**, e0230072 (2020).
21. Luies, L. & Loots, D. T. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics* **12**, 1–9 (2016).
22. Majors, R. E. *Sample preparation fundamentals for chromatography*. (Agilent Technologies, 2013).
23. Patti, G. J., Yanes, O. & Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell. Biol.* **13**, 263–269 (2012).
24. Schoeman, J. C., Du-Preez, I. & Loots, D. T. A comparison of four sputum pre-extraction preparation methods for identifying and characterising M. tuberculosis using GCxGC-TOFMS metabolomics. *J. Microbiol. Methods* **91**, 301–311 (2012).
25. Phiri, M. M. *Re-assessment and optimisation of an organic acid extraction method for automation* PhD thesis, North-West University (South Africa), Potchefstroom Campus, (2017).
26. Mouskeftara, T., Virgiliou, C., Theodoridis, G. & Gika, H. Analysis of urinary organic acids by gas chromatography tandem mass spectrometry method for metabolic profiling applications. *J. Chromatogr. A* **1658**, 462590. <https://doi.org/10.1016/j.chroma.2021.462590> (2021).
27. Pasikanti, K. K., Ho, P. C. & Chan, E. C. Development and validation of a gas chromatography/mass spectrometry metabolomic platform for the global profiling of urinary metabolites. *Rapid Commun. Mass Spectrom.* **22**, 2984–2992 (2008).
28. Blau, N., Duran, M., Gibson, K. M. & Vici, C. D. *Physician's Guide to the Diagnosis, Treatment, and Follow-Up of Inherited Metabolic Diseases*. Vol. 213 (Springer, 2014).
29. Fiehn, O., Kopka, J., Trethewey, R. N. & Willmitzer, L. Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry. *Anal. Chem.* **72**, 3573–3580 (2000).
30. Kanani, H., Chrysanthopoulos, P. K. & Klapa, M. I. Standardizing GC–MS metabolomics. *J. Chromatogr. B* **871**, 191–201 (2008).
31. Dunn, W. B. Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Phys Biol* **5**, 011001 (2008).
32. Shulaev, V. Metabolomics technology and bioinformatics. *Brief. Bioinform.* **7**, 128–139 (2006).
33. Nicholson, J., Connelly, J., Lindon, J. & Holmes, E. Metabonomics: a platform for studying drug toxicity and gene function. *Nat. Rev. Drug Discov.* **1**, 153–161 (2002).
34. Kruve, A., Leito, I. & Herodes, K. Combating matrix effects in LC/ESI/MS: The extrapolative dilution approach. *Analytica Chimica Acta* **651**, 75–80 (2009).
35. Mobley, H. L. Urease. *Helicobacter Pylori: Physiology and Genetics*, 177–191 (2001).
36. Qin, Y. & Cabral, J. M. Review properties and applications of urease. *Biocatal. Biotransform.* **20**, 1–14 (2002).
37. Kim, J., Choi, K. & Chung, D. in *Comprehensive Sampling and Sample Preparation* (ed Janusz Pawliszyn) 701–721 (Academic Press, 2012).
38. Taufik, M. *et al.* in *AIP Conference Proceedings. Sonication Technique for Nicotine Extraction From Saliva and Urine* (AIP Publishing, 2023).
39. Drabińska, N. *et al.* An optimization of liquid–liquid extraction of urinary volatile and semi-volatile compounds and its application for gas chromatography–mass spectrometry and proton nuclear magnetic resonance spectroscopy. *Molecules* **25**, 3651 (2020).
40. Gertsman, I. & Barshop, B. A. Promises and pitfalls of untargeted metabolomics. *J. Inher. Metabol. Dis.* **41**, 355–366 (2018).

Acknowledgements

This work is based on the research supported wholly/in part by the National Research Foundation (NRF) of South Africa (Grant no. 129871). The NRF provided funding for sample analysis for the entire larger study, of which this sub-study forms part of. All opinions, findings and conclusions, or recommendations expressed in this paper is that of the authors, and the NRF accepts no liability whatsoever in this regard.

Author contributions

C.O., B.A., and L.L. conceptualised the manuscript and drafted the paper (including all figures). C.O. and B.A. completed all experimental work. L.L. critically reviewed the content. All authors approved the final version to be submitted.

Funding

Open access funding provided by North-West University.

Competing interests

The authors declare that there are no competing interests, and that this manuscript, and the work described therein, is unpublished and has not been submitted for publication elsewhere.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44690-7>.

Correspondence and requests for materials should be addressed to L.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Supplementary information

For GC-MS analysis, the primary oven program is depicted in Figure S1. The initial temperature was 70°C, which was held for 1min. The temperature was then ramped up as follows: 5°C/min to 100°C, 10°C/min to 160°C, 13°C/min to 230°C, and finally 20°C/min to 300°C, which was held for 2 min. The secondary oven was programmed identically to that of the primary oven, except for a +5°C at each interval. The total run time of each sample was $\approx$ 24min.

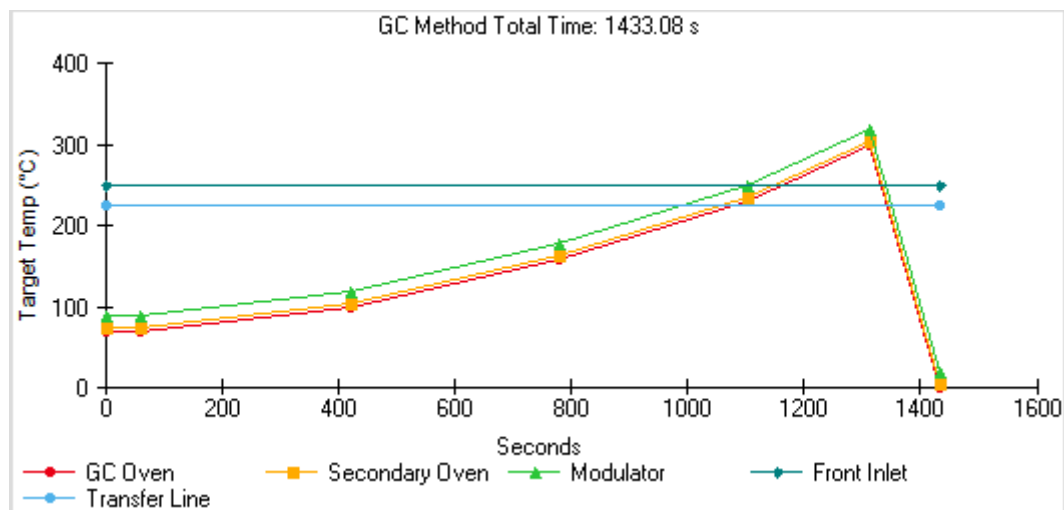


Figure S1: Temperature ramping during GC-MS analysis.

APPENDIX B

Olivier, C. and Luies, L. (2023). Optimising a urinary extraction method for non-targeted GC-MS metabolomics. Presented (as poster by L. Luies) at the 19th Annual Conference of the Metabolomics Society, Niagara Falls, Canada, 18–22 June 2023.

Optimising a urinary extraction method for non-targeted GC-MS metabolomics

Olivier, C. and **Luies, L.***

Human Metabolomics, North-West University, Private Bag X6001, Box 269, Potchefstroom, South Africa, 2531

*Presenter and correspondence: laneke.luies@nwu.ac.za

Introduction

Urine is ideal for non-targeted metabolomics, offering a broad representation of normal and pathological cellular processes for a holistic metabolic assessment. Gas chromatography mass spectrometry (GC-MS) is widely regarded as the "gold standard" for metabolomics studies¹ due to its numerous advantages. While organic acid (OA) extraction is commonly used for urinary GC-MS, it is better suited for targeted investigations in clinical settings. Considering the aim of non-targeted metabolomics to analyse various compound classes at varying concentrations, optimal extraction becomes crucial. Recent technological advancements have prompted the need to optimise the traditional OA extraction method or explore alternative approaches, such as a direct analysis (DA) method, which is quicker and more cost-effective. However, studies suggest that even the DA method could benefit from optimisation. For instance, additional drying between two-step derivatization² and/or urease sample pre-treatment to address challenges posed by urea in non-targeted urine metabolomics³. However, the latter remains a topic of debate within the scientific community due to its associated advantages and disadvantages.

Aim

To optimise a low-volume urine preparation procedure for non-targeted GC-MS analysis, covering various compound classes. Five methods were evaluated based on repeatability, metabolome coverage, and metabolite recovery. These methods included an OA extraction method (Method 1) and modified versions thereof such as the use of an additional extraction solvent (Method 2), two-step derivatization (Method 3), and a combination of Methods 2 and 3 (Method 4). Additionally, we tested a DA method (Method 5). The superior method was further optimised by investigating an additional drying step between two-step derivatization (Method 6) and urease sample pre-treatment (Method 7), while considering the effect of heating versus sonication.

Methods

Method 1

Traditional OA extraction⁴ (n=6)

Method 2

Additional extraction solvent⁵ (n=6)

Method 3

One-step vs two-step derivatization⁶ (n=6)

Method 4

Combination of Methods 2 & 3 (n=6)

Method 5

Direct analysis⁷ (n=6)

OA vs DA:

Evaluate validation parameters:

- Repeatability
- Metabolome coverage
- Metabolite recovery



Part A:

Which method is superior for non-targeted GC-MS urine analysis?



Method 6

- Group 1: Two-step derivatization (n=9)
- Group 2: Additional drying between two-step derivatization² (n=9)

Method 7

Exp. 1: Effect of urease pre-treatment³

- Group 1: Urease-treated; heating (n=6)
- Group 2: Urease-non-treated; heating (n=6)
- Group 3: Water-treated; heating (n=6)

Exp. 2: Heating vs sonication⁹

- Group 1: Urease-treated; heating (n=6)
- Group 2: Urease-non-treated; heating (n=6)
- Group 4: Urease-treated; sonication (n=6)
- Group 5: Urease-non-treated; sonication (n=6)



Part B:

Which method is most suited for future non-targeted GC-MS urine analysis?

Results and Discussion

Repeatability

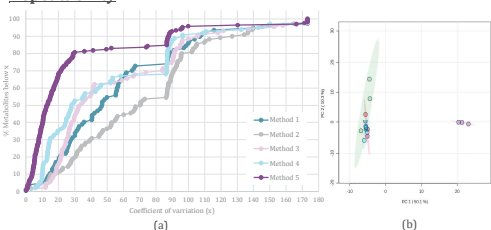


Figure 1: Repeatability of the five extraction methods. (a) The percentage metabolites below a given coefficient of variation (CV) threshold, ranked as Method 5 (82.4%), Method 3 (62.1%), Method 4 (61.9%), Method 1 (54.6%), and Method 2 (36.4%). (b) Principal component analysis (PCA) plot which provides an overview of the overall similarity or dissimilarity among the methods without directly incorporating the CV values.

Metabolome coverage

OA methods may not be suitable for all non-targeted metabolomics applications due to bias toward a specific compound class and limited compound recovery ability. In accordance, Method 5 achieved the highest metabolome coverage, with only 26.4% compounds remaining undetected, while detecting 91 unique metabolites from multiple compound classes. Next was Method 2 (51.4%; 13 unique), Method 4 (57.3%; 7 unique), Method 3 (61.8%; 1 unique), and Method 1 (66.4%; 0 unique).

Metabolite recovery

Based on organic, amino, fatty acids in a spiking solution: Method 5 recovered 57.7% of the compounds, followed by Method 2 (35%), Method 4 (28.3%), Method 1 (22.2%), and Method 3 (20.8%).

Overall performance

Table 1: Method rankings based on validation parameters, and an overall performance cumulative score of each.

	Repeatability	Metabolome coverage	Metabolite recovery	Cumulative score
Method 1 (original in-house)	4	5	4	13
Method 2 (acetonitrile addition)	5	2	2	9
Method 3 (two-step derivatization)	2	4	5	11
Method 4 (combination of Methods 2 & 3)	3	3	3	9
Method 5 (direct analysis)	1	1	1	3

Method 5 outperformed all other methods, consistently ranking first in every validation parameter.

Additional drying between two-step derivatization

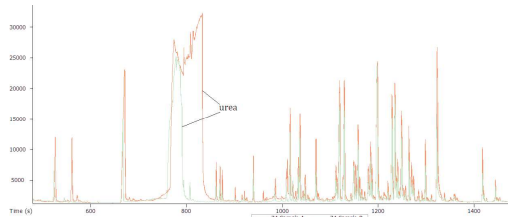


Figure 2: The total ion chromatograms of a sample subjected to two-step derivatization (Sample A; orange) and a sample subjected to an additional drying between two-step derivatization (Sample B; green). Sample B showed a significant reduction in the urea signal, which is considered advantageous.

Average peak height was visually better with conventional two-step derivatization, but the amino acid class displayed higher reported areas with additional drying². Inclusion of an additional drying step resulted in reduced signals of "junk compounds" (reagent artifacts).

Effect of urease sample pre-treatment

Table 2: Method rankings based on validation parameters, when comparing urease-treated, non-treated, and water-treated samples.

	Repeatability	Metabolome coverage	Metabolite recovery	Cumulative score
Urease-treated samples (UT-H)	3	3	3	9
Urease-non-treated samples (NT-H)	2	2	2	6
Water-treated samples (WT-H)	1	1	1	3

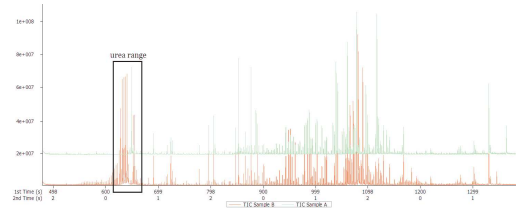


Figure 3: Total ion chromatograms of a urease-treated sample (Sample A; green) and a urease-non-treated sample (Sample B; orange). Urease pre-treatment showed a significant reduction in the signal of urea, which is considered advantageous, but also causes reduction in overall metabolic signals.

Heating vs sonication

Table 3: Method rankings based on validation parameters, when comparing heating and sonication following urease pre-treatment.

	Repeatability	Metabolome coverage	Metabolite recovery	Cumulative score
Urease-treated: Heating (UT-H)	4	4	4	12
Urease-treated: Sonication (UT-S)	2	3	2	7
Urease-non-treated: Heating (NT-H)	1	2	3	6
Urease-non-treated: Sonication (NT-S)	3	1	1	5

A urease non-treatment approach is better, regardless of heating or sonication.

Researchers should carefully assess the effects of urease pre-treatment and evaluate the trade-off between urea reduction and the potential compromise in metabolite recovery. However, with urease pre-treatment, sonication is more advantageous due to (i) efficient mixing and agitation; (ii) more homogeneous distribution of urease; (iii) better penetration into cellular structures; and (iv) avoidance of thermal degradation^{9,10}.

Conclusion

Our **low-volume DA method** have few analytical steps (which is more time efficient, cost-effective, and repeatable), and can extract much of the metabolome, making it ideal for non-targeted urine GC-MS analysis. This approach was further **improved by incorporating additional drying between two-step derivatization but did not benefit from urease sample pre-treatment**. Indeed, urease pre-treatment should be reserved for specific analyses targeting urea-related measurements, in which case sonication is recommended over heating.

References

1. Fiehn, O. Metabolomics by gas chromatography-mass spectrometry: Combined targeted and untargeted profiling. *Current protocols in molecular biology* 114, 30.34.1-30.34.32 (2016).
2. Liebecke, M. & Puskas, E. Drying enhances signal intensities for global GC-MS metabolomics. *Metabolites* 9, 68 (2019).
3. Webb-Robertson, B.-J. et al. A statistical analysis of the effects of urease pre-treatment on the measurement of the urinary metabolome by gas chromatography-mass spectrometry. *Metabolomics* 10, 897-908 (2014).
4. Luies, L. & Loots, D. T. Tuberculosis metabolomics reveals adaptations of man and microbe in order to outcompete and survive. *Metabolomics* 12, 1-9 (2016).
5. Majors, R. E. Sample preparation fundamentals for chromatography. (Agilent Technologies, 2013).

6. Patti, G. J., Yanes, O. & Siuzdak, G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* 13, 263-269 (2012).
7. Esterhuizen, K. et al. A urinary biosignature for mitochondrial myopathy, encephalopathy, lactic acidosis and stroke like episodes (MELAS). *Mitochondrion* 45, 38-45 (2019).
8. Palmas, F., Mussap, M. & Fattuoni, C. Urine metabolome analysis by gas chromatography-mass spectrometry (GC-MS): Standardization and optimization of protocols for urea removal and short-term sample storage. *Clinica Chimica Acta* 485, 236-242 (2018).
9. Taufik, M. et al. Sonication technique for nicotine extraction from saliva and urine. *AIP Conference Proceedings* 2626 (2023).
10. Kim, J., Choi, K. & Chung, D. S. In *Comprehensive Sampling and Sample Preparation* (ed Janusz Pawliszyn) 701-721 (Academic Press, 2012).

This work is based on the research supported wholly/in part by the National Research Foundation of South Africa (grant number 120871).

APPENDIX C

Olivier, C. and Luies, L. (2023). WHO Goals and Beyond: Managing HIV/TB co-infection in South Africa. SN Comprehensive Clinical Medicine, 5(1), p.251.
(<https://doi.org/10.1007/s42399-023-01568-z>)



REVIEW



WHO Goals and Beyond: Managing HIV/TB Co-infection in South Africa

Cara Olivier¹ · Laneke Luies¹

Accepted: 4 September 2023
© The Author(s) 2023

Abstract

The intricate relationship between HIV and TB, particularly in South Africa which grapples with an alarming prevalence of both diseases, presents a multifaceted challenge influenced by historical, social, and biological factors. This review explores the co-dependent nature of these diseases and the challenges to effective control strategies. The ramifications of Apartheid significantly shaped these pandemics, with TB thriving in impoverished conditions marked by overcrowding and limited health care access. Co-infection intensifies health challenges, further complicating treatment and management. Despite commendable efforts, achieving the ambitious “End TB” and “End HIV” goals set by the World Health Organization (WHO) for 2030 remains intricate due to socio-economic disparities, health care constraints, and political dynamics. The recent emergence of the COVID-19 pandemic introduced additional complexities, disrupting diagnostic and treatment services. Nevertheless, South Africa’s resilience is evident through destigmatisation campaigns, innovative preventative interventions, and significant contributions to global research. The importance of prevalence studies cannot be overstated; accurate data collection and analysis remain pivotal for informed interventions and tracking progress towards WHO goals. Here, we elucidate the need for a unified and comprehensive national response, supported by international collaboration, to effectively address the enduring HIV and TB burden within South Africa. As the nation navigates intricate historical legacies, social factors, and evolving health challenges, a crucial question arises: Will South Africa ultimately meet the WHO’s elimination goals or at least reduce the number of HIV-related TB deaths? Recognising potential obstacles, it becomes imperative to strategise responses to enable effective syndemic management and advancement towards WHO’s objectives.

Keywords TB · HIV · South Africa · Co-infection · WHO Goals · COVID-19

Introduction

The human immunodeficiency virus (HIV) and tuberculosis (TB) are two of the major infections that continue to be global health threats. Efforts have been undertaken to combat these diseases, with global reports highlighting the urgency of the issue. Global reports indicated there were 1.5 million new HIV infections and 680,000 deaths from acquired immune deficiency syndrome (AIDS)-related causes in 2021 alone. However, substantial challenges persist. An estimated 27.5

million people are currently on antiretroviral (ARV) medication globally. Of the 377 million people living with HIV, an estimated 10.2 million are not currently on HIV treatment [1]. Until 2019, TB was the leading cause of death from a single infectious agent, ranking above HIV/AIDS. While significant progress has been made, the number of patients newly diagnosed with TB between 2020 and mid-2021 in most of the high TB-burden countries was significantly below average compared to 2019. An immediate consequence of this can be observed in the number of TB deaths in 2020, estimated at 1.3 million people with an additional 214,000 deaths among HIV-positive people [2].

These two infections exert massive pressure on health care systems worldwide and even more so when combined due to an increase of resources needed to treat co-infected patients. The co-existence of these diseases presents unique difficulties, as HIV and TB have an interconnected relationship which causes severe adverse effects [2, 4]. Notably, HIV/TB co-infection prevalence is highest in

This article is part of the Topical Collection on *Medicine*

✉ Laneke Luies
laneke.luies@nwu.ac.za
Cara Olivier
caraolivier1@gmail.com

¹ Focus Area Human Metabolomics, North-West University, Potchefstroom Campus, Private Bag X6001, Box 269, Potchefstroom 2520, North-West, South Africa

the African region with southern Africa exceeding 50% of these cases. This indicates a substantial gap in testing and treatment for co-infected cases, as the amount of people diagnosed with TB that know their HIV status has decreased to 4.2 million [2].

To alleviate the disease burden, the World Health Organization (WHO) compiled a list of high-burden countries regarding global TB cases. While South Africa ranks 8th globally in terms of TB cases, accounting for 3.3% of the total global TB cases, it ranks 1st in HIV cases, with an estimated 7,800,000 people living with HIV—this accounts for 13% of the global HIV cases. In 2020, Africa accounted for a major 25% of the total global TB cases, with South Africa making up a substantial 13.2% of these. Africa was also reported as the region with the highest cases of HIV/TB co-infection (individuals infected with HIV and active TB), with South Africa contributing to an astounding 50% of these [1, 2]. Looking ahead to 2030, the WHO continues its resolute stance against TB through the “End TB Strategy”, which encompass a holistic strategy that includes (1) integrated patient-centred care to ensure equitable access for individuals affected by TB; (2) implementing inclusive policies and robust systems to address social determinants, strengthen health care systems, and provide financial protection; (3) driving innovation through intensified research; (4) mobilising communities for awareness and reducing stigma and discrimination; (5) fostering international collaborations for migrant populations; and (6) enhancing data-driven decision-making and adaptive strategies [2]. These interconnected elements collectively strive to eliminate the global burden of TB and promote a healthier world.

South Africa plays a pivotal role in advancing global efforts to combat TB by driving research, innovation in treatment, and facilitating clinical trials. This commitment extends to groundbreaking work on HIV preventative treatments, microbicide changes, and behavioural changes [5, 6]. The country has conducted several large-scale clinical trials and contributed research and resources to advance the development of an HIV vaccine [7]. Another standout achievement is the development of the Xpert MTB/RIF test, capable of detecting *Mtb* and rifampicin resistance [8]. Moreover, South Africa’s substantial genomic research on TB has unveiled genetic variations linked to increased TB susceptibility, providing important insights into its pathogenesis [9]. This dedication extends to TB vaccine research, exemplified by the South African Tuberculosis Vaccine Initiative (SATVI), which contributes to clinical trials worldwide [10].

The recent COVID-19 pandemic has introduced significant setbacks to the landscape of addressing HIV and TB in South Africa. The pandemic’s impact is multifaceted, leading to increased TB transmission within households due to isolation measures and instilling fear of seeking treatment outside homes, compounded by stigmatisation. This

disruption is further emphasised by the negative stigma associated with COVID-19 contraction [11]. Notably, the pandemic has disrupted essential TB diagnostic and treatment services, significantly impeding overall progress as the COVID-19 pandemic took precedence [2]. Consequently, the strides taken in reducing TB mortality rates globally have suffered a setback due to the pandemic’s influence [11].

Historical context has profoundly shaped South Africa’s experience of the HIV and TB pandemics, underscoring the need to grasp their socio-economic underpinnings to formulate effective strategies for case management and significantly reducing mortality in the twenty-first century. Considering the historical trajectory of HIV and TB in South Africa, the progress made in their mitigation, and the persistent challenges faced, a clearer perspective emerges on the feasibility of attaining the WHO and United Nations (UN) Sustainable Development goals [12].

Methods: Search Strategy and Selection Criteria

This article reports on findings of several literature reviews about HIV and TB in South Africa. It also reports on the findings of the HIV and TB reports of the WHO and UN which are published each year. This article focusses on the co-infection of HIV and TB and its current prevalence in South Africa. However, few articles have evaluated the syndemic between these two diseases in South Africa specifically, and therefore, relevant articles from international literature were used as indications of what we may find through future research also in the southern African countries. Clinical trials and research included in this review were selected based on the criteria of whether the trial yielded successful results and if it was recently published.

References for this review included published results from the WHO TB reports and The Joint United Nations Programme on HIV/AIDS (UNAIDS) from 1997 to 2022. Other relevant articles were identified through Google Scholar by using terms “HIV”, “TB”, “Co-infection”, and/or “South Africa”. Relevant articles resulting from these searches were considered and cited in this manuscript. Only articles published in English were included.

A Short History of HIV in South Africa

South Africa has a unique history regarding its response to the HIV/AIDS pandemic. Although some decisions have been scrutinised, it is important to understand what led to these decisions in order to avoid the same pitfalls when future crises or pandemics arise.

The first recorded AIDS death of a South African was in 1981 [13]. Although this was more than 40 years ago, the country is still experiencing the repercussions from the array of decisions that resulted in unnecessary loss of life. In the 1980s, it was a widespread global belief that HIV/AIDS was a homosexual disease, resulting in a majority of the population not taking the necessary precautions to prevent the spread of HIV [14]. The Apartheid government at the time did, however, little to contest these claims. There were several untrue labels placed on individuals that had AIDS following the start of this pandemic which included homosexuals, migrant mine workers, and even haemophilic patients. Between 1983 and 1985, several haemophilic patients acquired HIV by receiving contaminated blood, leading to the assumption that haemophilic patients were at risk. Another belief arose from a 1986 survey done on mine workers in South Africa. During the survey, it was found that some of the participants (0.01%) had HIV and that they were coincidentally all migrant workers. Following this survey, the government stated that immigrants were to blame for HIV, banning the entry of any HIV-infected migrant workers into the country [14]. Thus, the limited initial response to the pandemic can be attributed to the widespread beliefs that HIV/AIDS was a virus that only affected a specific type of person, namely homosexual men, haemophiliacs, and immigrant mine workers that entered the country [13, 14]. Since South Africa was ruled by Apartheid at the time, this misconception was advertised broadly to blame Black individuals and highlighted them as a threat to health, overall intensifying racial profiling [15]. This shifted the narrative to shaming the victims of the disease and blaming their infection on promiscuous behaviour [14].

In 1994, South Africa had its first democratic election during which Nelson Mandela was elected. Although this brought hope that the new government would dismantle racial inequalities and the stigma around HIV/AIDS, the pandemic still did not receive adequate resources and attention amidst all the new changes happening in the country [13]. In the following few years, the HIV mortality rate proceeded to surge [16] despite a very ambitious HIV/AIDS campaign that was outlined at the start of Nelson Mandela's presidency. Despite making headway, the programme ultimately fell short of expectations by 1998 [16].

The year 1999 was another turning point in the HIV pandemic with the election of Thabo Mbeki as president, who appointed Dr. Manto Tshabalala-Msimang as Minister of Health [17]. Tshabalala-Msimang questioned whether HIV caused AIDS and advocated beetroots, lemon, and garlic instead of ARVs for South Africans [18]. Based on advice received at the time, and the influence of other denialist advocates such as Peter Duesberg (a German–American molecular biologist and professor at Berkeley), South Africa's new president did not believe that HIV causes

AIDS but rather socio-economic factors such as poverty. He sent a letter to world leaders to convince them of his belief. The misinformation kept piling on with the government, spearheaded by Tshabalala-Msimang, which delayed ARV roll-out due to insinuations that they were ineffective and toxic. Eventually, in 2002, the South African high court ordered the production and roll-out of ARVs [17]. Nonetheless, Tshabalala-Msimang continued to spread misinformation and proceeded to promote alternative treatments—these were not scientifically proven and were merely based on nutrition and palliative care. It is no surprise that her unsupported claims and promotion of “African treatments” afforded her the name “Dr. Beetroot” and continues to be an embarrassing memory of South Africa's HIV past [19]. By 2006, mortality exceeded the number of new HIV infections, largely due to the slow roll-out of an antiretroviral therapy (ART) programme [14].

In 2009, Jacob Zuma was elected as president of South Africa. AIDS activists were sceptical about his appointment because of the controversial public statements [20] he made about HIV-positive women, i.e. that “taking a shower” would cut the risk of contracting HIV infection. However, the president left the denialistic claims of former president Thabo Mbeki behind and made progress in combatting the AIDS pandemic [17]. Since then, South Africa has made headway in achieving its goals to improve HIV mortality rates [1]. The current president, Cyril Ramaphosa, was elected in 2019 and has regarded the science behind the HIV/AIDS pandemic, which in turn allows him to make progress towards the goal to end the pandemic in South Africa [1, 21]. Although the current COVID-19 pandemic has taken a significant amount of South Africa's resources that could have been spent on the HIV pandemic, the future seems promising considering the proactive response of President Ramaphosa over the past 3 years of his term. This includes the call for increased treatment of HIV-positive individuals with ART and an overall positive, proactive approach towards HIV treatment [22, 23].

Upon evaluating the population percentage growth of South Africa in relation to the total HIV cases during the span of South Africa's presidencies from 1990 to 2021 (see supplementary material; Fig. S1), it becomes evident that despite the continuous population growth, the rise in HIV cases within the country has been disproportionate. The gap between the overall population of South Africa compared to the amount of HIV cases started to reduce between 1994 and 2002, indicating increased effort to combat the disease with ARV roll-outs as these roll-outs increased over the following years. A slight spike in HIV cases reported in 2020 is possibly linked to the COVID-19 pandemic. After 2020, it was expected that the amount of HIV cases would continue to increase. However, data from the WHO 2022 report indicates there was a decrease in the percentage of

total HIV cases from the previous year, with 2020 having a 4% increase from the previous year, and 2021 having a 3% increase from the previous year. Therefore, although the total amount of HIV cases in 2021 still increased from the previous year, the percentage increase was less comparatively. According to the UNAIDS 2022 update report, the global response towards the HIV crisis has been set back due to several global crises such as the COVID-19 pandemic that the world had to combat. Why then has the percentage of total HIV cases in South Africa not increased from the previous year?

The answer quite possibly lies in looking closer at the spike of HIV cases in 2020 (Fig. S1). In 2020, the percentage increase of HIV cases is the highest that it has been in 10 years. This disruption in the trend would suggest that the effects of the global COVID-19 pandemic were already visible in 2020, with the decrease in HIV cases shown in 2021 reflecting the increased knowledge and efforts in combating the COVID-19 pandemic [24], which had subsequently affected the HIV epidemic. The first roll-out of vaccinations for COVID-19 was administered at the beginning of 2021 in South Africa.

A Short History of TB in South Africa

Although TB has less political agendas attached to it, it is known as the “disease of poverty” in South Africa [25]. The WHO has published an annual global TB report since 1997 in which the prevalence, progress, and future goals towards the eradication of TB worldwide are discussed. As one of the countries with the highest amount of TB cases annually, South Africa is included in every report. Why does TB have such a foothold in the country, and what conditions in South Africa contribute to the high disease prevalence in poverty-stricken regions?

TB was first introduced to South Africa in the seventeenth century when European immigrants or colonists arrived in the country [14]. An interesting development led to the rapid spread of TB among the Black population of South Africa following this initial introduction. In the nineteenth century, South Africa began its explosive development of its mining industry. The conditions in the mines were grim with poor working conditions including being overcrowded, workers receiving low nutritious meals, and inhaling silica dust regularly, which contributed to the susceptibility and rapid spread of TB under these workers [14, 26]. It is proposed that the oscillatory movement of these mineworkers between their work and the rural communities that they lived in, which possessed minimal medical resources, majorly contributed to the spread of the disease [14].

In the twentieth century, TB cases rapidly and noticeably rose in South Africa, with a particular resurgence in the late 1990s [14]. During this period, the Apartheid government

spent the larger portion of the health care budget on areas where the White population resided although they were the minority. When analysing the per capita expenditure of different areas, the predominantly “White living areas” were given almost four times the amount of per capita expenditure for health care compared to “Black areas” [27]. The lack of adequate health care provided by the government only fuelled the spread of TB in a time when it was critical to face the threat of TB head-on. Arguably, by 1994, when the Apartheid government was abolished, the damage was already done, and TB posed a much greater threat to the country than initially anticipated.

In 1997, the National Tuberculosis Programme (NTP) was established to actively address the failing primary health care systems in place meant to deal with the TB endemic. One of the biggest milestones of this programme was the implementation of the directly observed treatment short-course (DOTS), which immensely strengthened the TB treatment programme [28]. Although this was a step in the right direction, the potential gain of TB control was less than expected, with major progress being countered by the maturing HIV pandemic which saw a rapid increase during this time [14]. However, the initiatives under the NTP continued ahead, not deterred by the apparent lack in progress.

In 1999, the WHO introduced DOTS-Plus, a programme launched to address the growing concern of multi-drug resistant (MDR) TB. This programme was aimed at building upon the five elements of DOTS (i.e., political commitment, microscopy services, drug supplies, surveillance and monitoring systems and use of highly efficacious regimens, and direct observation of treatment) while considering second-line anti-TB drugs, specifically the need to use them in resource-limited settings. The first pilot projects for DOTS-Plus were launched in 2000, and by 2006, these pilot projects showed significant evidence that demonstrated MDR TB can effectively be treated using DOTS-Plus [2]. The WHO’s TB guidelines were subsequently endorsed, and further treatment facilities were established for MDR TB by the NTP in South Africa. The DOTS-Plus programme was, however, only launched in South Africa in 2002 [29]. Since 2003, there were several more initiatives and milestones reached under the NTP. A discussion of these falls beyond the scope of this review, but the authors recommend the thorough review of Churchyard et al. [28], should the reader be interested.

Despite the introduction of the MDR TB guidelines, MDR TB and extensively drug-resistant (XDR) TB completely took over the nation, increasing the national emergency that TB posed. In 2007, a study was done that identified XDR TB cases in all nine provinces of South Africa, highlighting the extreme prevalence of drug-resistant TB strains. Both the DOTS and ART programmes are threatened by these drug-resistant TB strains due to the exclusion

of first-line TB medication when patients are resistant to isoniazid and rifampicin [14, 30], the main drugs used for treating drug-susceptible TB. Second-line TB medication, used for treating drug-resistant TB, is not ideal as these medications place a huge burden on the liver, and side-effects often include liver inflammation, jaundice, joint pain, and indigestion. These side-effects in turn may ultimately lead to patient non-adherence [31, 32].

Since 2007, progress has been made towards reducing the number of TB deaths in South Africa, although certain setbacks were encountered [2]. South Africa's first-ever prevalence study was conducted in 2018. This was a major advancement as it enabled the WHO to conduct more accurate global estimates and to provide important estimates of the burden of TB in South Africa. It also brought new information on health-seeking behaviour of patients to light [33].

When regarding the history of both HIV and TB, although they are two separate diseases of viral and bacterial descent, respectively, they have a highly interconnected relationship, posing a deadly threat to high-burden countries like South Africa. The control of these pandemics is fundamental to the future economic growth and political progress of the country [34]. The historical conditions, specifically the impact of Apartheid, have shaped and defined the nature of HIV and TB pandemics in South Africa as we combat these diseases in the twenty-first century. Although major progress and headway have been made since the first TB case in South Africa, high TB case prevalence and mortality rates continue to be a concern in the country. If the end of the HIV and TB pandemics is to be achieved, a better national response is required from not only the country's health minister but also a unified effort between researchers and health care professionals alike.

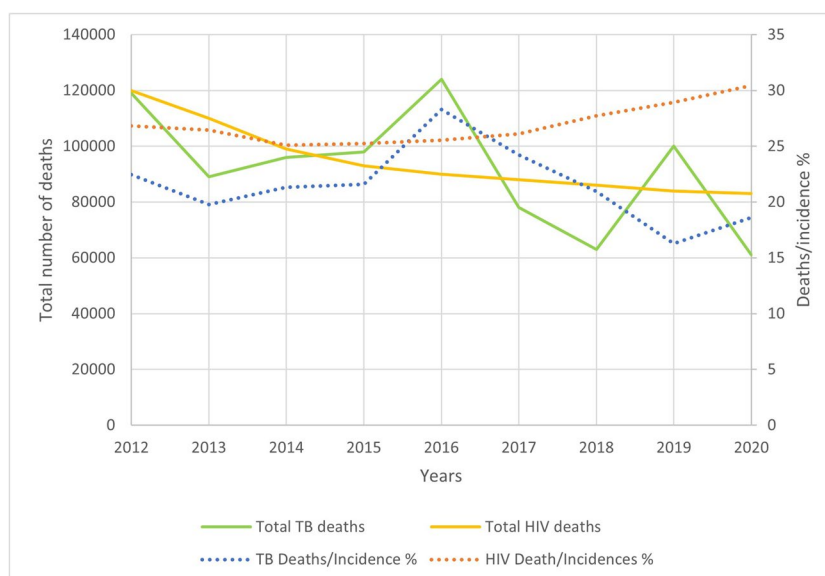
When comparing the difference between the number of TB deaths in HIV-negative people compared to those in HIV-positive people in South Africa (Fig. S2), the detrimental effect of HIV/TB co-infection is evident. HIV and TB attenuate each other as the number of TB deaths in HIV-positive people far exceeds the number of TB deaths in HIV-negative people. Therefore, it is evident that that relationship between HIV and TB should be investigated further.

The Co-dependent Relationship Between HIV and TB

Africa has the highest rate of HIV/TB co-infection in the world, with a large contribution stemming from South Africa. When two diseases combine, they form an even deadlier syndemic, referring to exacerbated health consequences of the biological interactions observed in multiple individuals within a population when both are present [35]. The deadly consequences related to simultaneous infection of a host have been referred to as “the cursed duet”.

In newly HIV-diagnosed patients, TB is the most common presenting illness and the most prevalent opportunistic disease among people living with HIV [36]. It was reported that among the total worldwide incident cases of TB, 8% of these were people living with HIV. In Africa, the proportion of TB cases co-infected with HIV was the highest as reported in 2020. Not only are HIV patients more susceptible to contracting TB, but it also lowers treatment success rates. The success rate among co-infected patients is 77% compared to 86% (for first-line TB treatment regimens) in those with only TB [2]. Contributing to the low treatment outcomes, is the rapid development of drug-resistant strains which complicates treatment [36]. The comparison between HIV and TB deaths in South Africa can be seen in Fig. 1. This figure depicts the relationship between the amount of

Fig. 1 The relationship between the total deaths to deaths divided by incidence percentage, for HIV and TB in South Africa. Total TB deaths are indicated with a green line, TB deaths vs incidence percentage is indicated with a dotted blue line, total HIV deaths are indicated with a yellow line, and the HIV deaths vs incidence percentage is indicated with a dotted orange line



HIV and TB deaths divided by their subsequent incidences expressed as a percentage. This percentage illustrates the correlation between the mortality rate and the incidence of each disease, revealing the number of individuals who succumb to the infection. This percentage has remained relatively steady and experienced a significant decrease since 2020. An important contributing factor to note is that proper treatment for HIV can substantially extend individuals' lives, which can contribute to this figure [2].

Deaths due to both HIV and TB have been decreasing from around 120,000 in 2012 to around 83,000 for HIV (orange line in Fig. 1) and 61,000 for TB (green line in Fig. 1) in 2020. It is, therefore, clear that the prevention of TB deaths has been more successful than the prevention of HIV deaths over the last 8 years in South Africa. This is illustrated by the death divided by incidence percentage dotted lines in the graph. TB incidence has decreased to around 18% in 2020, whereas HIV incidence has been rising steadily to around 30% in 2020. Thus, in the case of HIV infections, preventative treatment is relatively successful whereas treatment of the AIDS patients is not that successful. This can also be seen in the number of HIV infections per 1,000 people that has been reduced from 8.4% in 2012 to 4.6% in 2020.

The amount of total HIV and TB deaths in South Africa has decreased from 2020 to 2021. Taking the COVID-19 pandemic into account, it can be argued that since HIV-positive individuals' immune systems are more vulnerable, these individuals have a higher chance of contracting COVID-19 and subsequently their death was instead recorded as such and not exclusively an HIV death. The data and records of the causes of death might thus be affected by a global health threat such as the COVID-19 pandemic.

But what causes HIV and TB to have such an intricate relationship and how do they potentiate each other? Although we do not understand the detailed mechanisms of this interaction [4], HIV and TB enhance each other causing increased and accelerated deterioration of immunological functions, which, when treated poorly, causes an early death for the patient. HIV has been identified in high-burden countries, such as South Africa, as the most important risk factor to develop active TB, and in patients with latent TB, it increases the susceptibility to primary infection or the risk of TB reactivation [37]. As a patient's CD4 + T-cell count decreases, their probability to progress to active TB increases [36]. This increased susceptibility of HIV-infected patients to TB can greatly be attributed to two main mechanisms, namely (1) the increased reactivation of latent TB, or (2) the increased risk of contracting *Mtb*. The possibility of both situations being present suggests that these disease states should be closely monitored [2]. There are several features of patients with HIV that make them more susceptible to TB infection. One of the main features of HIV/AIDS is

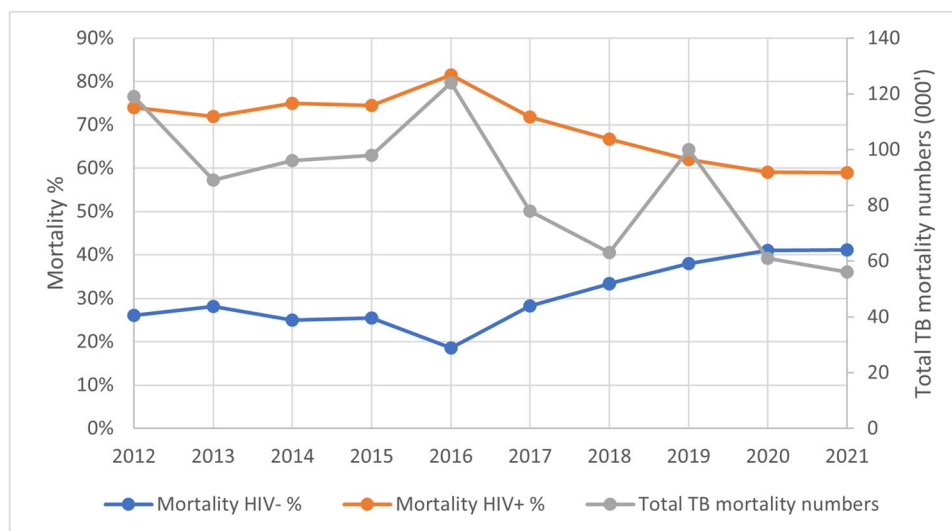
the immunosuppression that causes a loss of CD4 + T-cells. However, the underlying mechanism of this increased susceptibility is clearly more complicated as it has been shown that TB susceptibility increases soon after HIV infection, which means the CD4 + T-cell count is not the only culprit [2, 38]. Another feature of HIV/AIDS is the sustained immune response induced by the host since a prolonged immune activation leads to an enhanced disease progression, therefore making these patients more susceptible to TB [38]. These mechanisms are complex and poorly understood, necessitating further research to enhance comprehension for improved diagnostics and treatment approaches in the HIV/TB co-infected disease state, including the extent to which disease attenuation occurs and the associated metabolic alterations [4, 39]. However, conventional methodologies employing labour-intensive biochemical assays [39] have demonstrated alterations in net protein balance [40], hypoalbuminemia [41], and similar body composition shifts [42] in co-infected individuals relative to individuals with HIV infection or TB alone. In 2023, a pioneering study explored the serum metabolic profiles of healthy controls, untreated TB-positive patients, untreated HIV/TB co-infected patients, and HIV/TB co-infected patients on ART. This investigation unveiled substantial shifts in lipid and protein metabolism, alongside translocation of microbial products from the gut to the bloodstream. The findings imply a synergistic amplification of TB by HIV, contributing to increased inflammation, oxidative stress, ART-induced mitochondrial damage, and detrimental effects on gut health, thereby influencing energy availability. Although ART partially counteracts these trends in co-infected patients, further research is warranted to fully comprehend these changes and their implications [43].

From 2012 to 2020, total TB mortality in South Africa has been showing an overall declining trend (grey line in Fig. 2). From 2016 to 2020, the TB death percentage based on the total TB deaths reported has been declining for HIV-positive individuals with TB (orange line) compared to HIV-negative individuals (blue line), which could indicate that HIV has been better managed in South Africa since 2016.

Social and Biological Dynamics in South Africa that Contribute to HIV/TB Co-infection

The intertwined histories of both HIV and TB in South Africa have significantly influenced the syndemic challenges in the country today. However, specific social and biological dynamics play pivotal roles in perpetuating the spread of these diseases despite advancements in medical technology. TB has often been named as a “social disease” or “disease of poverty”, alluding to the living conditions and communal behaviours that contribute to its propagation [44]. Given the substantial proportion of South Africans living in poverty [45], conditions conducive to HIV and TB transmission

Fig. 2 TB mortality as a percentage of total TB mortality for HIV-positive (orange line) and HIV-negative (blue line) subjects, with total TB mortality numbers (grey line), shown from 2012 to 2021



prevail. Several studies underscore overcrowding as a notable driver of increased TB transmission [46, 47]. Close-knit living in informal settlements escalates the risk of both HIV and TB dissemination [48]. Furthermore, prevailing social challenges include inadequate food security, limited health care access, substandard living conditions, and a lack of disease awareness [38, 44]. The absence of proper education and treatment engagement in these communities exacerbates disease propagation. The intertwining of impoverished circumstances cultivates an environment that promotes HIV and TB transmission [38]. Moreover, the dual burden of social stigmatisation associated with HIV and TB, which can manifest as fear of discrimination or isolation, as well as misconceptions about the diseases, impedes individuals in communities from seeking essential treatment [49]. Effectively addressing disease management may necessitate an integral focus on rectifying the social and biological contexts within South African communities, a vital step towards significant progress.

Current Progress Supporting South Africa's Efforts Against HIV and TB

South Africa has recognised the significance of social dynamics in the battle against HIV and TB. Numerous campaigns have been undertaken to confront the stigma and hurdles encountered by communities when seeking treatment. An example of such an effort was initiated by the South African National AIDS Council (SANAC), aiming to counteract persistent social perceptions that continues to fuel stigma and discrimination against people living with HIV. The campaign featured the slogan "I can't change my HIV status but I can change my attitude" [50]. Subsequent to this, multiple other campaigns have been introduced in South Africa to educate and destigmatise these diseases within communities.

South Africa has played a significant role in global HIV research. Notably, it has contributed to HIV preventative treatments, including pre-exposure prophylaxis (PrEP). A recent study in South Africa examined the initiation and adherence to oral PrEP in a group of young women, providing crucial insights on expanding HIV prevention options for this vulnerable group in a high-burden HIV country like South Africa [51]. Further research has also focused on microbicide and behavioural changes [5, 6]. South Africa has been a site for several large-scale clinical trials of HIV prevention methods with a total of 21 clinical trials initiated up until 2020 [52], and even more being initiated since then. Notable successful trials include the HPTN 052 trial, which demonstrated the remarkable efficacy of ART in safeguarding against sexual transmission of the HIV virus. This ground-breaking study revealed that initiating ART in HIV-infected individuals before severe immune system compromise led to a substantial 93% reduction in HIV transmission. These findings underscore the enduring and dependable protective benefits of ART in preventing the spread of the virus through sexual contact [53]. Another impactful clinical trial was the ASPIRE trial, a phase III randomised, double-blind, placebo-controlled study evaluating the efficacy of the dapivirine vaginal ring. This study was conducted across 15 sub-Saharan African sites, including South Africa [54]. South Africa also contributed extensively to HIV treatment research over the years, exploring the impact and adherence of South African cohorts to ART administration [55]. South Africa has taken crucial steps to identify important barriers relating to treatment regimes, proposing several strategies aimed at enhancing treatment administration in the country to improve outcomes [56–58]. Efforts towards HIV vaccine development are being pursued through the South African AIDS Vaccine Initiative (SAAVI), established in 1999 in collaboration with government, ESKOM, and the Medical

Research Council [59]. South Africa has a high prevalence of both HIV and TB, and co-infection research, considering their metabolic impact, is being thoroughly researched in the country. A recent review examined the complex metabolic effects of HIV and TB and proposed several future objectives [4]. While a recent HIV vaccine trial, Mosaico phase III, failed in preventing HIV [60], the fight against HIV in South Africa continues unabated. Promising trials for the M72/AS01E TB vaccine are ongoing in South Africa in 2023 [7], and an innovative approach using allogeneic haematopoietic stem cell transplantation has shown success in curing HIV, with the involvement of Honorary Professor, Dr. Monique Nijhuis [61].

In combating TB, South Africa has made significant contributions, both scientifically and financially. Notable milestones include the development of the Xpert MTB/RIF test, which can detect *Mtb* and rifampicin resistance [8]. South Africa participated in a multi-country study evaluating the diagnostic accuracy of this test, collaborating with the Foundation for Innovative New Diagnostics (FIND) [8, 62]. The active participation of the country in this study, along with several other studies conducted by South African researchers, has played a vital role in the successful roll-out of this test. Notably, Professor Nicol and fellow researchers ascertained the accuracy of the test for diagnosing pulmonary TB in a South African Hospital [63]. Numerous clinical trials conducted in South Africa have contributed to TB research, including a landmark clinical trial by the Aurum Institute that demonstrated the efficacy of a shorter, simpler TB regime [64]. Significant genomic research on TB has also been conducted in South Africa, identifying genetic variations that increase susceptibility to TB and providing important insights into the pathogenesis of the disease [9]. The country has shown immense dedication to TB vaccine research through SATVI, contributing to clinical trials worldwide [10]. Notably, in December 2022, the WHO implemented a new all-oral 6-month treatment regimen for drug-resistant TB in South Africa, showing promising results [65]. The Deputy President of South Africa has launched a National Strategic Plan (NSP) to combat HIV and TB, emphasising breaking down barriers, ensuring equitable and equal access to HIV and TB services, building resilient health care systems, and establishing revitalised, inclusive, and accountable institutions [66]. This plan also acknowledges the impact of pandemics, such as the COVID-19 pandemic, on the fight against HIV and TB, highlighting the need to address burdens caused by such outbreaks. However, it is important to address underlying social economic challenges before the country can achieve the newly outlined goals of the NSP [67–69].

While South Africa is actively striving to combat both HIV and TB, the numerous clinical trials and research

initiatives, while crucial in the fight against these diseases, do not adequately address the unique social barriers within South Africa that impede progress. The NSP acknowledges these barriers and delineates a well-structured strategy for achieving its four primary goals. However, despite the commendable efforts, there are areas where the practical implementation of these strategies faces challenges. South Africa's history of economic challenges, corruption, and political turmoil presents significant obstacles to achieving the NSP's goals [70]. The South African health sector exhibits a range of corrupt practices, including bribery, theft, and subpar service quality [71]. International research has indicated that in underdeveloped countries, corrupt practices are experienced by up to 80% of the population within the health sector [72]. With South Africa's national elections coming up in 2024, the successful implementation of the NSP by the current Deputy President may encounter further obstacles in the coming years.

Adding Fuel to the Fire

It is evident that the syndemic of HIV and TB has been a major burden for South Africa for a very long time, which was worsened with the recent COVID-19 pandemic. One needs to consider the specific aspects that have contributed to the COVID-19 pandemic adding fuel to the fire, as well as their influence on future progress towards ending HIV and TB.

The WHO 2021 global report estimated that 1.4 million fewer people received care for TB in 2020, a 21% reduction compared to the previous year. The decrease is proposed to be caused by COVID-19-related disruptions such as lockdown implementation and the subsequent difficulty in delivering medicine and administering treatment with restricted access to clinics, hospitals, or treatment centres. Reduced accesses for TB diagnosis and treatment have been one of the biggest effects of COVID-19, ultimately leading to increased TB deaths. TB case detection and reporting have also been negatively affected as COVID-19 case detection and treatment have taken precedence. Several specific disruptions also impacted TB detection, specifically the overall health system capacity being reduced, reduced ability of people seeking health care services during national lockdowns being put in place, fear of contracting the COVID-19 virus during clinic visits, or stigmas associated with symptoms that are similar between TB and COVID-19, making people less willing to report their symptoms [2].

The HIV pandemic, specifically in South Africa, was also severely impacted by the COVID-19 pandemic. There was a 50% decrease in HIV testing and ART initiation at the beginning of the COVID-19 lockdown, which only marginally

increased over the following months [73]. Different aspects of the COVID-19 pandemic have contributed negatively towards the progress of ending HIV. Dorward et al. [73] suggested a paucity of personal protective equipment, space for physical distancing in clinics, reduced opening times of clinics, and staff being redeployed from HIV testing and care to COVID-19 response activities played a crucial role in deterring HIV care. These additional challenges posed by COVID-19 added stress to a country that also faces severe poverty and issues with overall access to clinics in rural communities. A major challenge for South Africa currently is to reverse the effects COVID-19 had on ending both HIV and TB. The future WHO and UNAIDS reports might shed more light on the impact that COVID-19 had on these two diseases; however, the preliminary data already hints at major setbacks in both the fight against HIV and TB.

When considering South African's TB deaths relative to the reported incidences, and comparing it to that of HIV, for the period of 2012 to 2021, a predicted future relationship can be determined (Fig. S3). If the current trends are followed, the TB number will be higher than expected in 2022/2023, which might reflect the impact of COVID-19 and the overall increase in death experienced by the country. It might, however, still be too soon to see the effects of COVID-19 on HIV as the WHO predicts that 2022 figures will reflect this effect better.

Prevalence Studies and Reliability of Data

Prevalence studies give an accurate measurement of disease burden in a population. These studies are of great importance to governments, health professionals, policymakers, and the public sector of a country as they report on the development of diseases and the success of health care delivered in an attempt to stop the disease from spreading [74]. Before 2020, there had not been any public data available reporting on TB prevalence in South Africa, despite the WHO strongly urging South Africa to conduct an internal prevalence study since 2008. Thus, the data that the WHO uses to compile annual estimates for South Africa is greatly based on modelling exercises. This means that the reported estimates from the WHO are unreliable as the data is based on clinical reports rather than actual data provided by the country [33]. To understand the true burden of TB in South Africa, and provide realistic estimates, a prevalence study was finally initiated in 2018 and published in 2021 [75]. The results showed that 737 new TB cases were reported in South Africa, in 2018, which is significantly higher than the number of new cases reported by the WHO, which was 520 cases [75, 76]. Limitations of the prevalence study worth mentioning include (i) the performance of the Xpert Ultra tests used for diagnosing patients, as various false positive results were produced and the data was subsequently

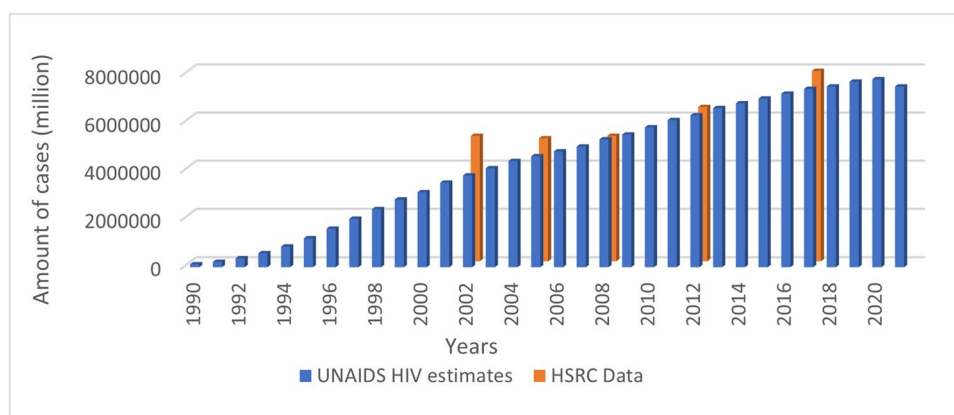
adjusted and conservatively amended to account for this; (ii) the participation rate of the study was lower than expected with an overall participation of only 66.1%; (iii) the HIV status of the participants was based on self-reporting; and (iv) only certain participants were eligible for sputum examination [75]. Nonetheless, the overall impact of this study is overwhelmingly positive, as it provides valuable data that the WHO can use to improve their reporting on the TB prevalence in South Africa as well as supplying important cohort information between sub-groups such as the prevalence of TB in male and female participants of different ages. This prevalence study also sets a good precedence for future TB studies in South Africa.

The Human Sciences Research Council (HSRC), however, started conducting HIV prevalence studies in South Africa in 2002, followed by another five in 2005, 2008, 2012, 2017, and 2022. These studies are done in the form of national household surveys and are essential for supplying accurate HIV statistics for South Africa. However, this type of survey has drawbacks such as the exclusion of certain groups, e.g. children younger than 2 years old who may have been infected through mother-to-child transmission, as well as persons living in institutions like prisons, military barracks, and boarding schools. Of note is the much higher HIV prevalence in South Africa determined by the HSRC compared to early estimates by UNAIDS, as reported in 2018, which can be seen in Fig. 3. The more recent estimates made by UNAIDS in regard to HIV prevalence in South Africa is, however, more in line with HSRC statistics implying that UNAIDS have considered and adjusted their estimates in recent years to be more accurate [77]. Entities that report on global data, such as the WHO and UNAIDS, have certain criteria for data collection from various countries that are used to compile important estimates for annual reports. Reporting agencies and other sources such as hospitals and clinics within countries submit their yearly data, which is then analysed using specific software programmes. Since a country's entire population cannot be tested for HIV or TB, only a portion of the population is selected, and estimates are made based on the selected data produced by the group. When countries conduct their own prevalence studies and population-based surveys as previously mentioned, these estimations can be adjusted to realistic data for accuracy, which aids in ending these pandemics.

Are the "End TB" and "End HIV" Goals Out of Reach?

As mentioned previously, the WHO has set targets for ending TB. One of the main goals is to reduce the absolute number of TB deaths by 90% by the year 2030. The number of TB deaths (including HIV-positive and HIV-negative cases) in South Africa has not been declining linearly over

Fig. 3 Total HIV cases in South Africa as reported by the UNAIDS (blue bars) compared to the total HIV cases in South Africa as reported by the HSRC (orange bars), which is a South African internal study



the years (Fig. S4). Following the latest data from 2021, if the TB cases are to decline with 90% by 2030, a strengthened control approach is urgently needed to reach the 2030 goal. Similarly, the number of AIDS-related deaths among the South African population has been declining steadily over the years. If the same goals are to be applied to the number of AIDS-related deaths in South Africa, the goal would be a 90% reduction in the death toll by the year 2030. Considering current trends (Fig. S4 and Fig. S5), achieving the 2030 goal is highly unlikely—unless an immense, strengthened effort is made and maintained.

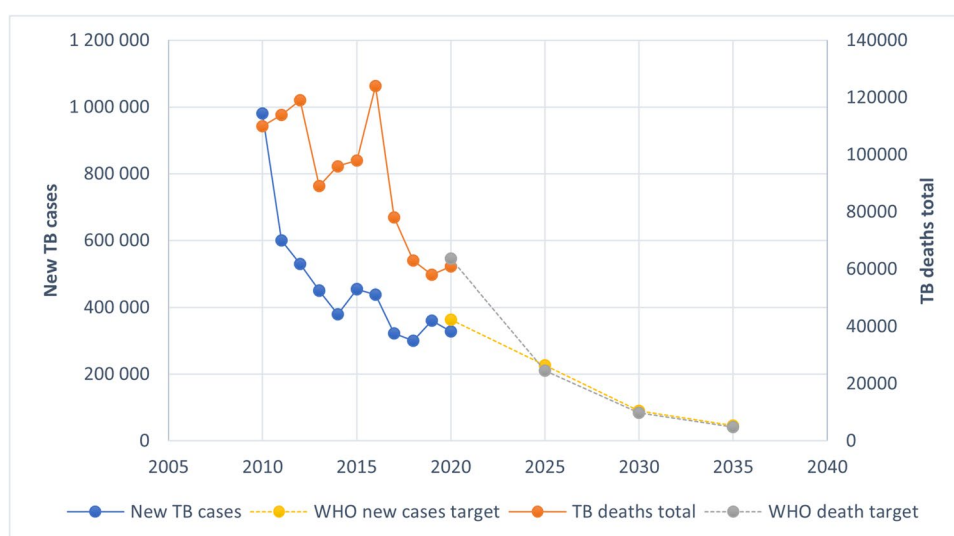
The WHO also provides more detailed goals to encourage the fight against TB. Figure 4 illustrates the TB death reduction goals set by the WHO (represented by the grey dots) for 2020 (35% reduction), 2025 (75% reduction), 2030 (90% reduction), and 2035 (95% reduction). Although South Africa seems to be on the right track in reducing TB mortality, the country will have to strengthen its overall effort in combatting TB to achieve the WHO goals.

The goals set by the WHO are meant to motivate countries to end the TB pandemic. As previously mentioned,

the WHO's "End TB Strategy" for 2030 involves a comprehensive approach, including patient-centred care, inclusive policies, innovative research, community mobilisation, global collaborations for migrants, and data-driven adaptive strategies to combat TB. Although these goals are usually accompanied by several suggested strategies on how to achieve them [2], they are not personalised to specific countries. Since each country faces its own set of challenges, social dynamics and stigmas should be addressed first before the ultimate goal of ending TB can be achieved.

Although South Africa has made progress in the reduction of TB incidences with the country's research and NSP implementation, these might not be enough to achieve the "End TB" goals. When considering similar efforts made by the country between 2015 and 2020, the goals were not met. The aim was to reduce TB incidences by 20%, and only 9% of the incidences were reduced by 2020. The death rate was also aimed to be reduced by 35%, and only a 14% reduction was achieved by 2020 [78]. It is thus evident that the long-term "End TB" goals will not be achievable if the problems at hand faced by South Africa are not dealt with first.

Fig. 4 New TB cases (blue line) and total deaths (orange line) in South Africa, including the World Health Organization's global targets (grey line) for 2020, 2025, 2030, and 2035. The new case targets are indicated with a yellow line



Conclusion

The complex interplay of history, society, and health crises has moulded the landscape of HIV and TB in South Africa, impacting the country's battle against these stubborn pandemics. Apartheid's shadow still looms, contributing to a disproportionate burden on vulnerable groups and a complex co-infection dynamic. Although progress has been achieved through research and interventions, the persistent high disease rates highlight the urgency of the challenge. The recent COVID-19 pandemic has further complicated matters, undoing previous gains and revealing health care system vulnerabilities. Yet, South Africa's commitment to accurate data via prevalence studies offers hope for informed strategies. Achieving the ambitious "End TB" and "End HIV" goals demands more than data; it requires a coordinated approach, strong health systems, fair access, and robust institutions. While the journey ahead is challenging, past accomplishments driven by resilience demonstrate progress is possible. Confronting historical injustices, overcoming stigmas, and advancing co-infection research require collective effort. Through shared determination, South Africa's pursuit of a healthier, fairer future reflects its resilience and dedication to its people's well-being.

Abbreviations *AIDS*: Acquired immune deficiency syndrome; *ART*: Antiretroviral therapy; *ARV*: Antiretroviral; *COVID-19*: Coronavirus disease 2019; *DOTS*: Directly observed treatment short-course; *FIND*: Foundation for Innovative New Diagnostics; *HIV*: Human immunodeficiency virus; *HSRC*: Human Sciences Research Council; *MDR*: Multi-drug resistant; *Mtb*: Mycobacterium tuberculosis; *NTP*: National Tuberculosis Programme; *PrEP*: Pre-exposure prophylaxis; *SAAVI*: South African AIDS Vaccine Initiative; *SATVI*: South African Tuberculosis Vaccine Initiative; *TB*: Tuberculosis; *UNAIDS*: Joint United Nations Programme on HIV/AIDS; *UNAIDS*: The Joint United Nations Programme on HIV/AIDS; *WHO*: World Health Organization; *XDR*: Extensively drug-resistant

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42399-023-01568-z>.

Author Contribution CO and LL conceptualised the manuscript. CO designed/constructed the figures. LL and CO evaluated current literature for drafting the paper, followed by critical review. All authors approved the final version to be submitted.

Funding Open access funding provided by North-West University. This work is based wholly/in part on the research supported by the National Research Foundation of South Africa (grant no. 129871).

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Joint United Nations Program on HIV/AIDS (UNAIDS). Report on the global AIDS epidemic 2021. Geneva, Switzerland: UNAIDS. 2021.
2. World Health Organization. Global tuberculosis report. Geneva: World Health Organization; 2021.
3. Bruchfeld J, Correia-Neves M, Källenius G. Tuberculosis and HIV coinfection. *Cold Spring Harb Perspect Med*. 2015;5(7):a017871.
4. Liebenberg C, Luies L, Williams AA. Metabolomics as a tool to investigate HIV/TB co-infection. *Front Mol Biosci*. 2021;8:1–19.
5. Weinrib R, Browne E, Shapley-Quinn M, van der Straten A, Bek-sinska M, Mgodini N, et al. Perspectives from young South African and Zimbabwean women on attributes of four (placebo) vaginal microbicide delivery forms. *AIDS Behav*. 2020;24:637–47.
6. Karim SSA, Baxter C. Microbicides & their implications in HIV prevention. *Indian J Med Res*. 2010;132(6):656.
7. Harris RC, Quaife M, Weerasuriya C, Gomez GB, Sumner T, Boz-zani F, et al. Cost-effectiveness of routine adolescent vaccination with an M72/AS01E-like tuberculosis vaccine in South Africa and India. *Nat Commun*. 2022;13(1):602.
8. Lawn SD, Nicol MP. Xpert® MTB/RIF assay: development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. *Future Microbiol*. 2011;6(9):1067–82.
9. Iorger TR, Koo S, No E-G, Chen X, Larsen MH, Jacobs WR Jr, et al. Genome analysis of multi-and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa. *PLoS One*. 2009;4(11):e7778.
10. Hussey G. Prospects for new tuberculosis vaccines. *S Afr Med J*. 2007;97(10):1001–2.
11. Visca D, Ong C, Tiberi S, Centis R, D'ambrosio L, Chen B, et al. Tuberculosis and COVID-19 interaction: a review of biological, clinical and public health effects. *Pulmonology*. 2021;27(2):151–65.
12. United Nations: sustainable development goals. <https://sdgs.un.org/goals> Accessed 2023.
13. Simelela N, Venter W, Pillay Y, Barron P. A political and social history of HIV in South Africa. *Curr HIV/AIDS Rep*. 2015;12(2):256–61.
14. Karim SSA, Churchyard GJ, Karim QA, Lawn SD. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *Lancet*. 2009;374(9693):921–33.
15. Van der Vliet V. AIDS: losing "the new struggle"? *Daedalus*. 2001;130(1):151–84.
16. Abdool-Karim Q, Abdool-Karim SS. The evolving HIV epidemic in South Africa. *Int J Epidemiol*. 2002;48.

17. McNeil J. A history of official government HIV/AIDS policy in South Africa. In: South African History Online (SAHO). 2015. <https://www.sahistory.org.za/article/history-official-government-hivaids-policy-south-africa>. Accessed 2020.
18. Wetzel D. HIV positive: saints, sinners, and AIDS protest movements in Fig' Queer Religion. The Scholar and Feminist Online. 2011. <https://sfonline.barnard.edu/hiv-positive-saints-sinners-and-aids-protest-movements-in-fig-trees-queer-religion/>. Accessed 2021.
19. Baleta A. "Dr Beetroot" sidelined in South Africa. *Lancet Infect Dis*. 2006;6(10):620.
20. Robins S. Sexual politics and the Zuma rape trial. *J South Afr Stud*. 2008;34(2):411–27.
21. Nordling L. Our epidemic could exceed a million cases'—South Africa's top coronavirus adviser. *Nature*. 2020;583(7818):672.
22. Harris J. Confronting legacies and charting a new course? The politics of coronavirus response in South Africa. *Coronavirus Polit: Comp Polit Policy of COVID-19*. 2021:580–99.
23. Sidley P. South Africa's sex workers will receive HIV prevention and treatment. *BMJ*. 2016;352:i1593.
24. Naidoo V. Mobilising the public sector to combat COVID-19, and the pandemic's effect on public sector governance. In: Fouire P, Lamb G, editors. *The South African Responses to COVID-19: The Early Years*. Routledge; 2023. p. 21.
25. Andersson N. Tuberculosis and social stratification in South Africa. *Int J Health Serv*. 1990;20(1):141–65.
26. Packard RM. White plague, black labor: Tuberculosis and the political economy of health and disease in South Africa. Univ of California Press; 1989. p. 23.
27. Price M. Health care as an instrument of apartheid policy in South Africa. *Health Policy Plan*. 1986;1(2):158–70.
28. Churchyard G, Mametja L, Mvusi L, Ndjeka N, Pillay Y, Heselting A, et al. Tuberculosis control in South Africa: successes, challenges and recommendations: tuberculosis control-progress towards the Millennium Development Goals. *S Afr Med J*. 2014;104(3):244–8.
29. Pooran A, Pieterse E, Davids M, Theron G, Dheda K. What is the cost of diagnosis and management of drug resistant tuberculosis in South Africa? *PLoS One*. 2013;8(1):e54587.
30. Seung KJ, Keshavjee S, Rich ML. Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harb perspect Med*. 2015;5(9).
31. De Villiers L, Du Loots T. Using metabolomics for elucidating the mechanisms related to tuberculosis treatment failure. *Curr Metabolomics*. 2013;1(4):306–17.
32. Awofeso N. Anti-tuberculosis medication side-effects constitute major factor for poor adherence to tuberculosis treatment. *Bull World Health Organ*. 2008;86: B-D.
33. Pillay Y, Mvusi L, Mametja L, Dlamini S. What did we learn from South Africa's first-ever tuberculosis prevalence survey? *SAMJ*. 2021;111(4).
34. Whiteside A. Demography and economics of HIV/AIDS. *Br Med Bull*. 2001;58(1):73–88.
35. Singer M. A dose of drugs, a touch of violence, a case of AIDS, part 2: further conceptualizing the SAVA syndemic. *Free Inquiry in Creative Sociology*. 2006;34(1):39–54.
36. Tiberi S, Carvalho ACC, Sulis G, Vaghela D, Rendon A, Mello FCDQ, et al. The cursed duet today: tuberculosis and HIV-coinfection. *La Presse Médicale*. 2017;46(2):e23–39.
37. Kwan CK, Ernst JD. HIV and tuberculosis: a deadly human syndemic. *Clin Microbiol Rev*. 2011;24(2):351–76.
38. Badri M, Ehrlich R, Wood R, Pulerwitz T, Maartens G. Association between tuberculosis and HIV disease progression in a high tuberculosis prevalence area. *Int J Tuberc Lung Dis*. 2001;5(3):225–32.
39. Liebenberg C, Luies L, Williams AA. Metabolomics as a tool to investigate HIV/TB co-infection. *Front Mol Biosci*. 2021;8:692823.
40. Paton N, Ng Y, Chee C, Persaud C, Jackson A. Effects of tuberculosis and HIV infection on whole-body protein metabolism during feeding, measured by the [15N] glycine method. *Am J Clin Nutr*. 2003;78:319–25. <https://doi.org/10.1093/ajcn/78.2.319>.
41. Van Lettow M, Fawzi WW, Semba RD. Triple trouble: the role of malnutrition in tuberculosis and human immunodeficiency virus co-infection. *Nutr Rev*. 2003;61(3):81–90. <https://doi.org/10.1301/nr.2003.marr.81-90>.
42. Paton NI, Ng Y-M. Body composition studies in patients with wasting associated with tuberculosis. *Nutrition*. 2006;22(3):245–51. <https://doi.org/10.1016/j.nut.2005.06.009>.
43. Herbert C, Luies L, Loots DT, Williams AA. The metabolic consequences of HIV/TB co-infection. *BMC Infect Dis*. 2023;23(1):536.
44. Lönnroth K, Jaramillo E, Williams BG, Dye C, Ravignione M. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Soc Sci Med*. 2009;68(12):2240–6.
45. Mdululi P, Dunga S. Determinants of poverty in South Africa using the 2018 general household survey data. *J Poverty*. 2022;26(3):197–213.
46. Baker M, Das D, Venugopal K, Howden-Chapman P. Tuberculosis associated with household crowding in a developed country. *J Epidemiol Community Health*. 2008;62(8):715–21.
47. Pelissari DM, Diaz-Quijano FA. Household crowding as a potential mediator of socioeconomic determinants of tuberculosis incidence in Brazil. *PLoS One*. 2017;12(4):e0176116.
48. Lienhardt C. From exposure to disease: the role of environmental factors in susceptibility to and development of tuberculosis. *Epidemiol Rev*. 2001;23(2):288–301.
49. Mawar N, Sahay S, Pandit A, Mahajan U. The third phase of HIV pandemic: social consequences of HIV/AIDS stigma & discrimination & future needs. *Indian J Med Res*. 2005;122(6):471.
50. Peltzer K, Parker W, Mabaso M, Makonko E, Zuma K, and Ramlagan S. Impact of national HIV and AIDS communication campaigns in South Africa to reduce HIV risk behaviour. *Sci World J*. 2012.
51. Mansoor LE, Lewis L, Naicker CL, Harkoo I, Dawood H, Naidoo K, et al. Prospective study of oral pre-exposure prophylaxis initiation and adherence among young women in KwaZulu-Natal, South Africa. *J Int AIDS Soc*. 2022;25(7):e25957.
52. Laher F, Bekker L-G, Garrett N, Lazarus EM, Gray GE. Review of preventative HIV vaccine clinical trials in South Africa. *Adv Virol*. 2020;165:2439–52.
53. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. *N Engl J Med*. 2016;375(9):830–9.
54. Garcia M, Luecke E, Mayo AJ, Scheckter R, Ndase P, Kiweewa FM, et al. Impact and experience of participant engagement activities in supporting dapivirine ring use among participants enrolled in the phase III MTN-020/ASPIRE study. *BMC Public Health*. 2021;21:1–11.
55. Moosa A, Gengiah TN, Lewis L, Naidoo K. Long-term adherence to antiretroviral therapy in a South African adult patient cohort: a retrospective study. *BMC Infect Dis*. 2019;19(1):1–12.
56. Maskew M, Bor J, MacLeod W, Carmona S, Sherman GG, Fox MP. Adolescent HIV treatment in South Africa's national HIV programme: a retrospective cohort study. *Lancet HIV*. 2019;6(11):e760–8.
57. Coetzee D, Hildebrand K, Boule A, Maartens G, Louis F, Labatula V, et al. Outcomes after two years of providing antiretroviral treatment in Khayelitsha. *South Africa Aids*. 2004;18(6):887–95.

58. Miller CM, Ketlhapile M, Rybasack-Smith H, Rosen S. Why are antiretroviral treatment patients lost to follow-up? A qualitative study from South Africa. *Tropical Med Int Health*. 2010;15:48–54.
59. Tucker TJ, Mazithulela G. Development of an AIDS vaccine: perspective from the South African AIDS Vaccine Initiative. *BMJ*. 2004;329(7463):454–6.
60. Venkatesan P. HIV vaccine trial failure. *Lancet Infect Dis*. 2023;23(4):410.
61. Jensen BEO, Knops E, Cords L, Lübke N, Salgado M, Busman-Sahay K, Estes JD, Huyveneers LE, Perdomo-Celis F, Wittner M, Gálvez C. In-depth virological and immunological characterization of HIV-1 cure after CCR5Δ32/Δ32 allogeneic hematopoietic stem cell transplantation. *Nat med*. 2023;29(3):583–7.
62. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med*. 2010;363(11):1005–15.
63. Nicol MP, Workman L, Isaacs W, Munro J, Black F, Eley B, et al. 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 Infect Dis*. 2011;11(11):819–24.
64. Salazar-Austin N, Dowdy DW, Chaisson RE, Golub JE. Seventy years of tuberculosis prevention: efficacy, effectiveness, toxicity, durability, and duration. *Am J Epidemiol*. 2019;188(12):2078–85.
65. Organization WH: WHO announces landmark changes in treatment of drug-resistant tuberculosis. <https://www.who.int/news/item/15-12-2022-who-announces-landmark-changes-in-treatment-of-drug-resistant-tuberculosis> (2022). Accessed.
66. Mlangeni P: South Africa one of 30 high TB-burdened countries. <https://www.sabcnews.com/sabcnews/south-africa-one-of-30-high-tb-burdened-countries/> (2023). Accessed.
67. DeSanto D, Velen K, Lessells R, Makgopa S, Gumede D, Fielding K, et al. A qualitative exploration into the presence of TB stigmatization across three districts in South Africa. *BMC Public Health*. 2023;23(1):1–10.
68. Foster N, Vassall A, Cleary S, Cunnam L, Churchyard G, Sinanovic E. The economic burden of TB diagnosis and treatment in South Africa. *Soc Sci Med*. 2015;130:42–50.
69. Loveday M, Zweigenthal V. TB and HIV integration: obstacles and possible solutions to implementation in South Africa. *Tropical Med Int Health*. 2011;16(4):431–8.
70. Schneider H, Blaauw D, Gilson L, Chabikuli N, Goudge J. Health systems and access to antiretroviral drugs for HIV in Southern Africa: service delivery and human resources challenges. *Reprod Health Matters*. 2006;14(27):12–23.
71. Mantzaris E, Pillay P. Legal profession and corruption in health care: some reflective realities in South Africa. *Front Public Health*. 2021;9:688049.
72. Holmberg S, Rothstein B. Dying of corruption. *Health Econ Policy Law*. 2011;6(4):529–47.
73. Dorward J, Khubone T, Gate K, Ngobese H, Sookrajh Y, Mkhize S, et al. The impact of the COVID-19 lockdown on HIV care in 65 South African primary care clinics: an interrupted time series analysis. *The Lancet HIV*. 2021;8(3):e158–65.
74. Munn Z, Moola S, Riitano D, Lisy K. The development of a critical appraisal tool for use in systematic reviews addressing questions of prevalence. *Int J Health Policy Manag*. 2014;3(3):123.
75. Van der Walt M, Moyo S. The first national TB prevalence survey South Africa. 2018. <https://repository.hsra.ac.za/handle/20.500.11910/21854>. Accessed 2018.
76. World Health Organization. Global progress report on HIV, viral hepatitis and sexually transmitted infections, 2021: Accountability for the global health sector strategies 2016–2021: actions for impact. 2021.
77. Human Sciences Research Council. South African National HIV Prevalence, Incidence, Behaviour and Communication Survey. 2017. <https://hsra.ac.za/uploads/pageContent/10779/SABSSM%20V.pdf>. Accessed 2018.
78. Chakaya J, Khan M, Ntoumi F, Aklillu E, Fatima R, Mwaba P, et al. Global tuberculosis report 2020—reflections on the global TB burden, treatment and prevention efforts. *Int J Infect Dis*. 2021;113:S7–12.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Material for WHO Goals and Beyond: Managing HIV/TB Co-infection in South Africa

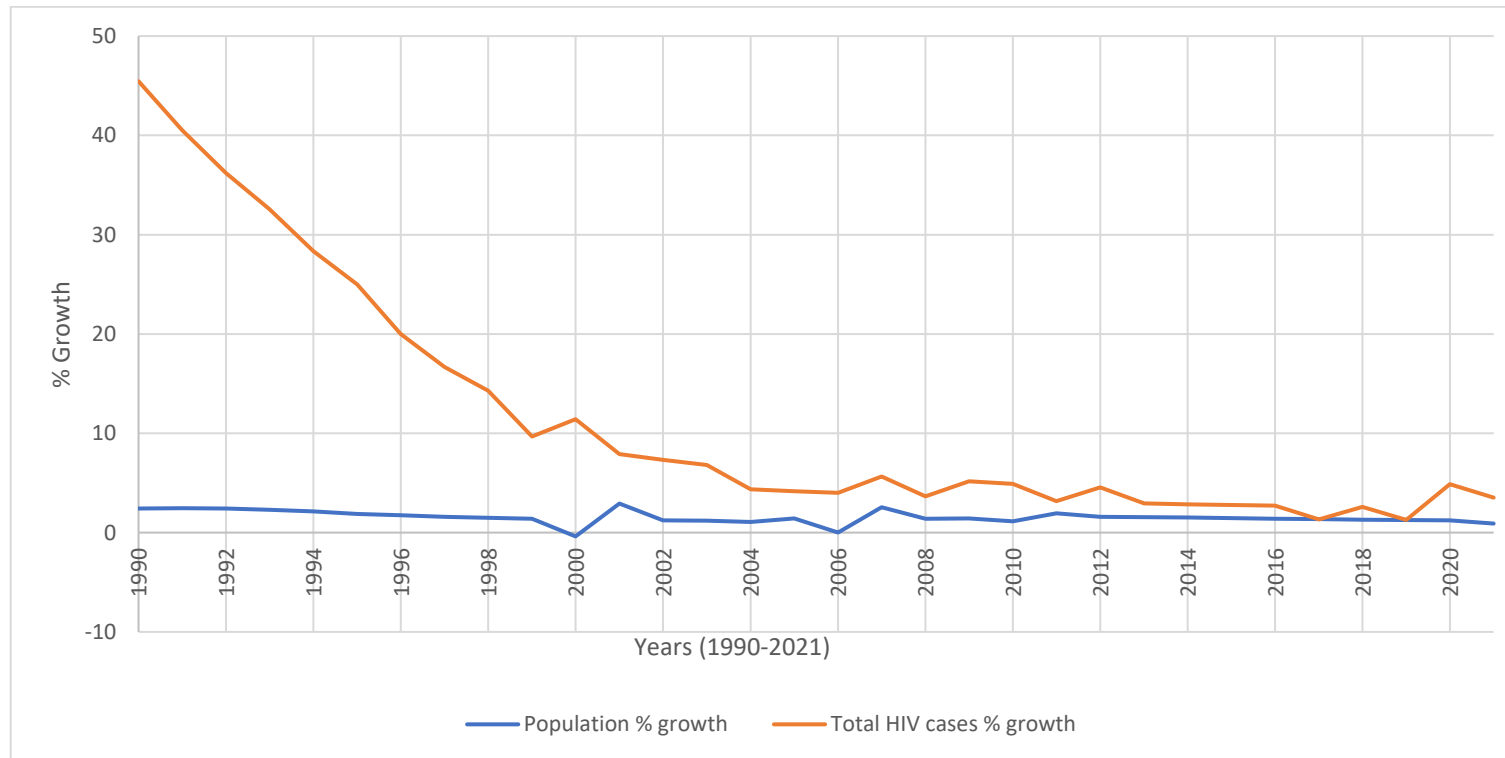


Fig. S1: South African population percentage (%) growth (blue line) compared to South African total HIV cases percentage growth (orange line), as shown from 1990 to 2021

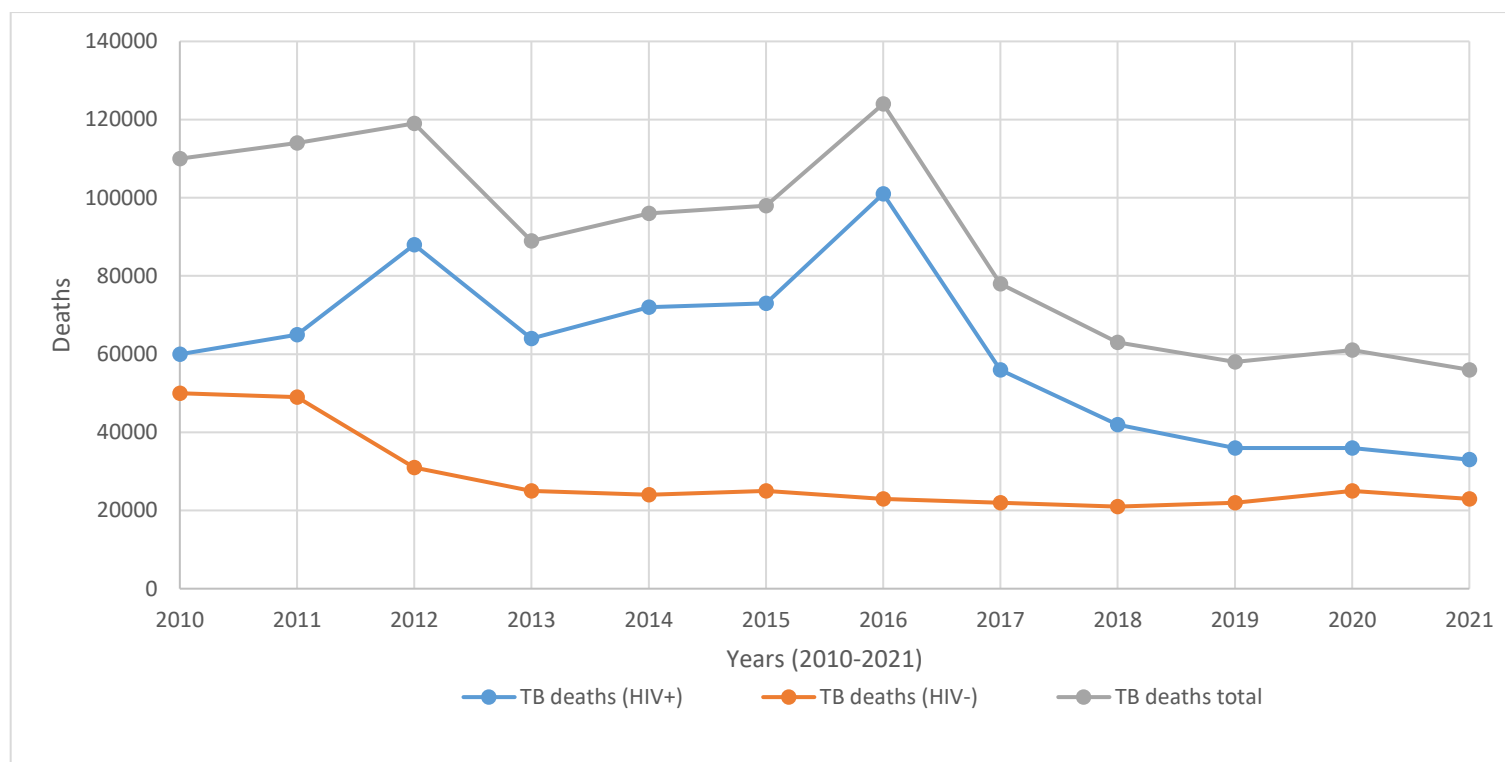


Fig. S2: TB deaths in HIV-positive people (blue line) compared to TB deaths in HIV-negative people (orange line) in South Africa from 2010 to 2021. The total number of TB deaths is also indicated (grey line)

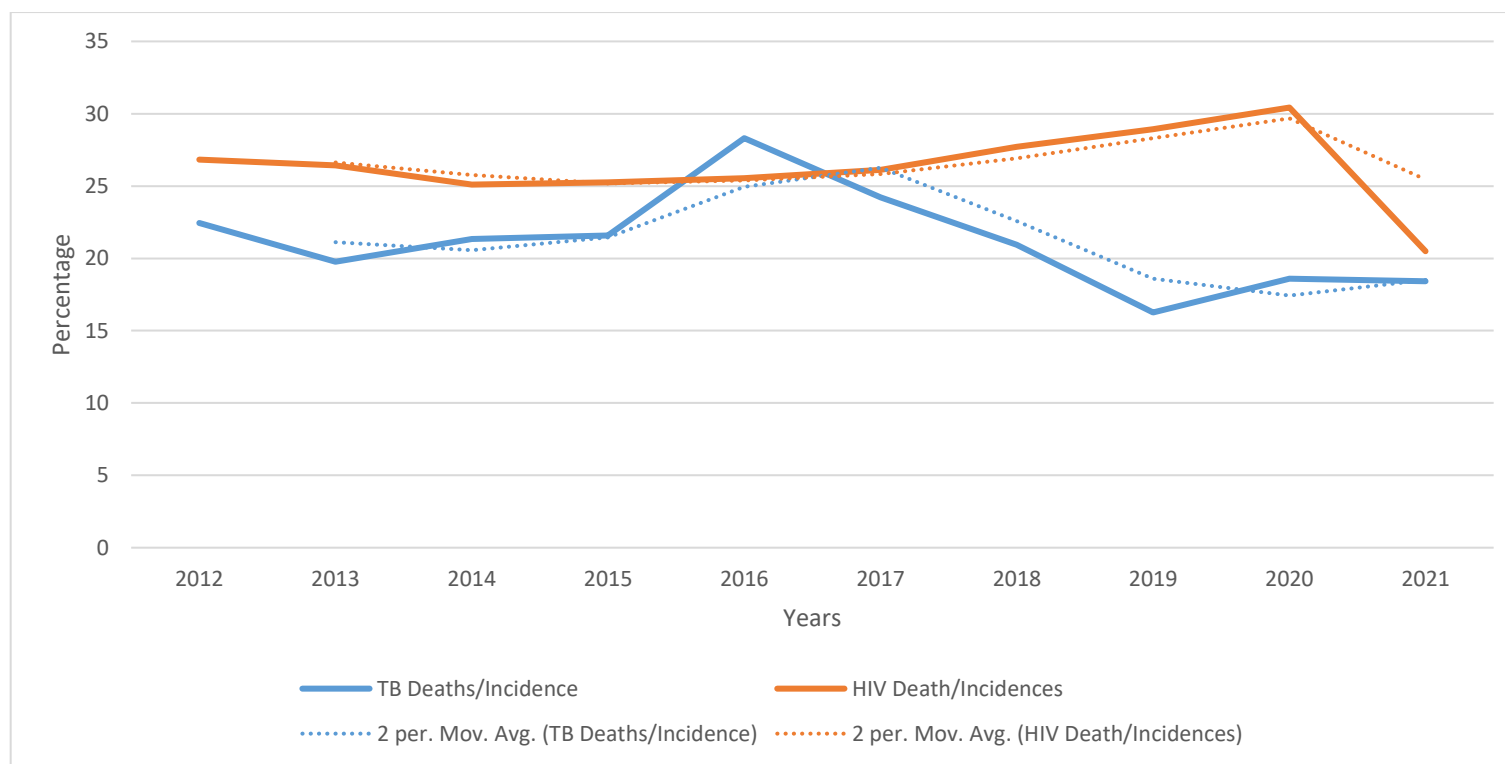


Fig. S3: South African TB deaths divided by incidences expressed as a percentage compared to the South African HIV deaths divided by incidences expressed as a percentage (blue and orange solid lines represent TB and HIV respectively). A moving average is included as a dotted line (blue and orange for TB and HIV respectively). Abbreviations: “2 per. Mov. Avg.”: two period moving average

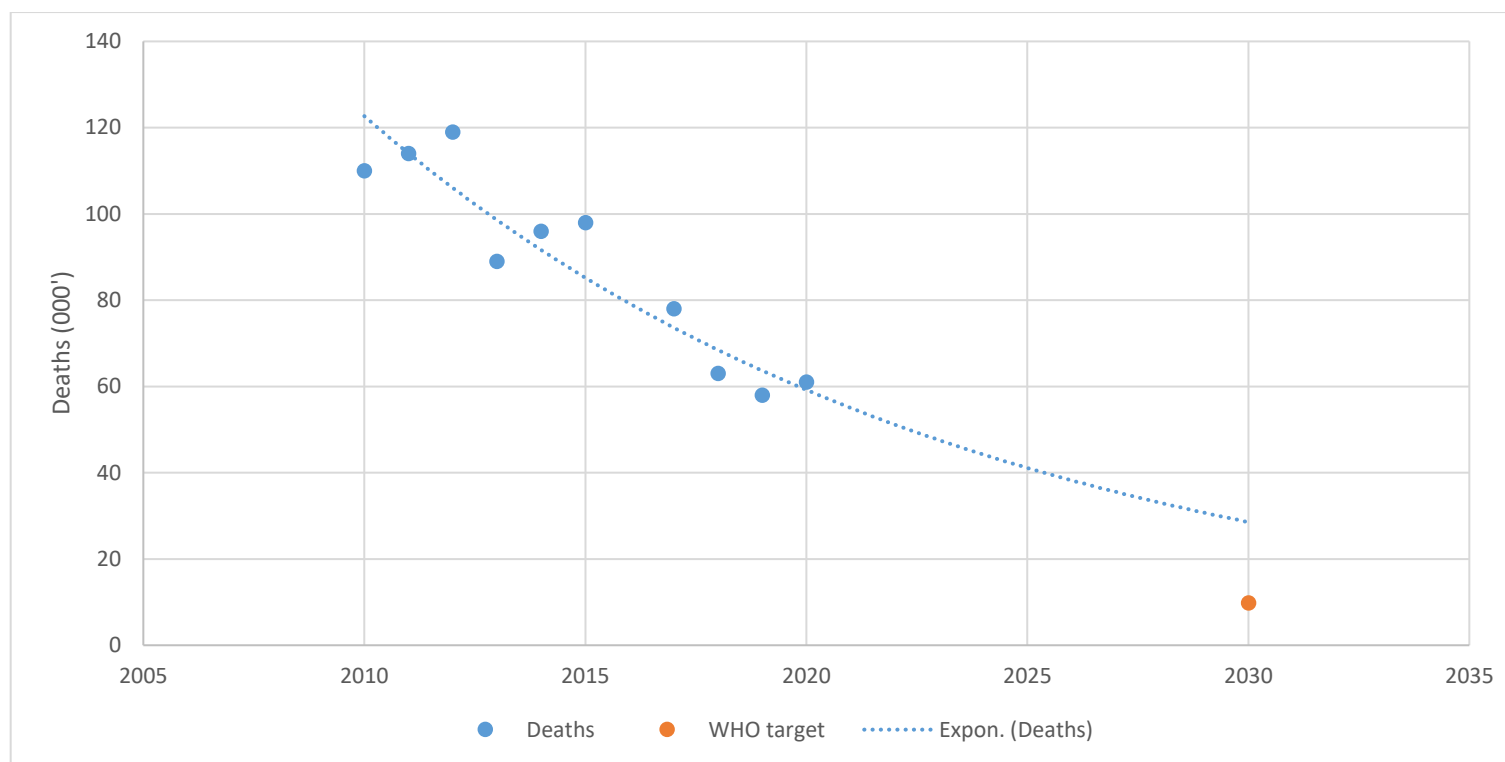


Fig. S4: The number of TB deaths (blue line) in South Africa from 2010 to 2020, with the World Health Organization's goal for 2030 represented by the orange dot. An exponential trend line was added (the dotted blue line) to depict a forward projection. Note that the death estimate for 2016 was removed as an outlier as it was disproportionately high with no apparent reason for this high estimate by the World Health Organization

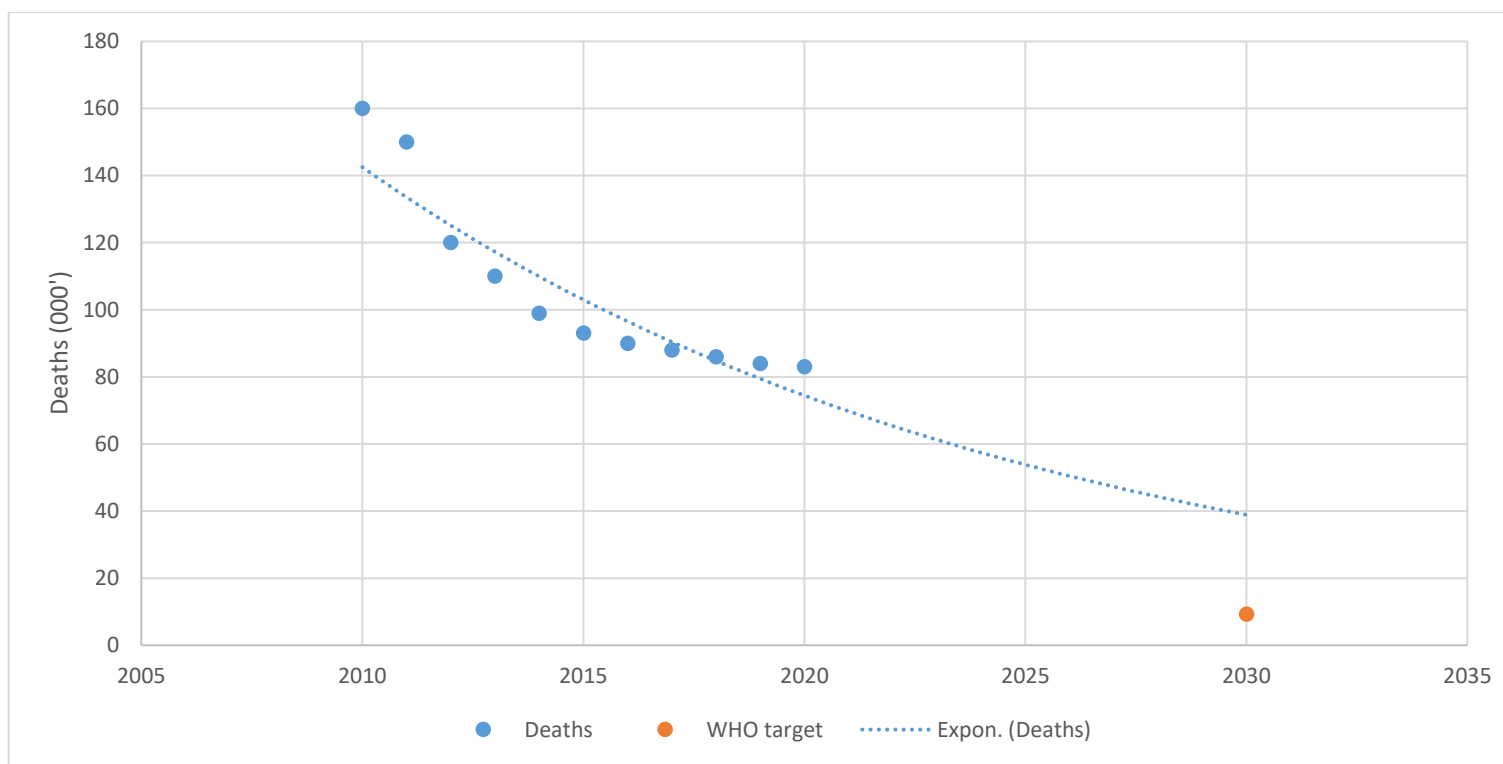
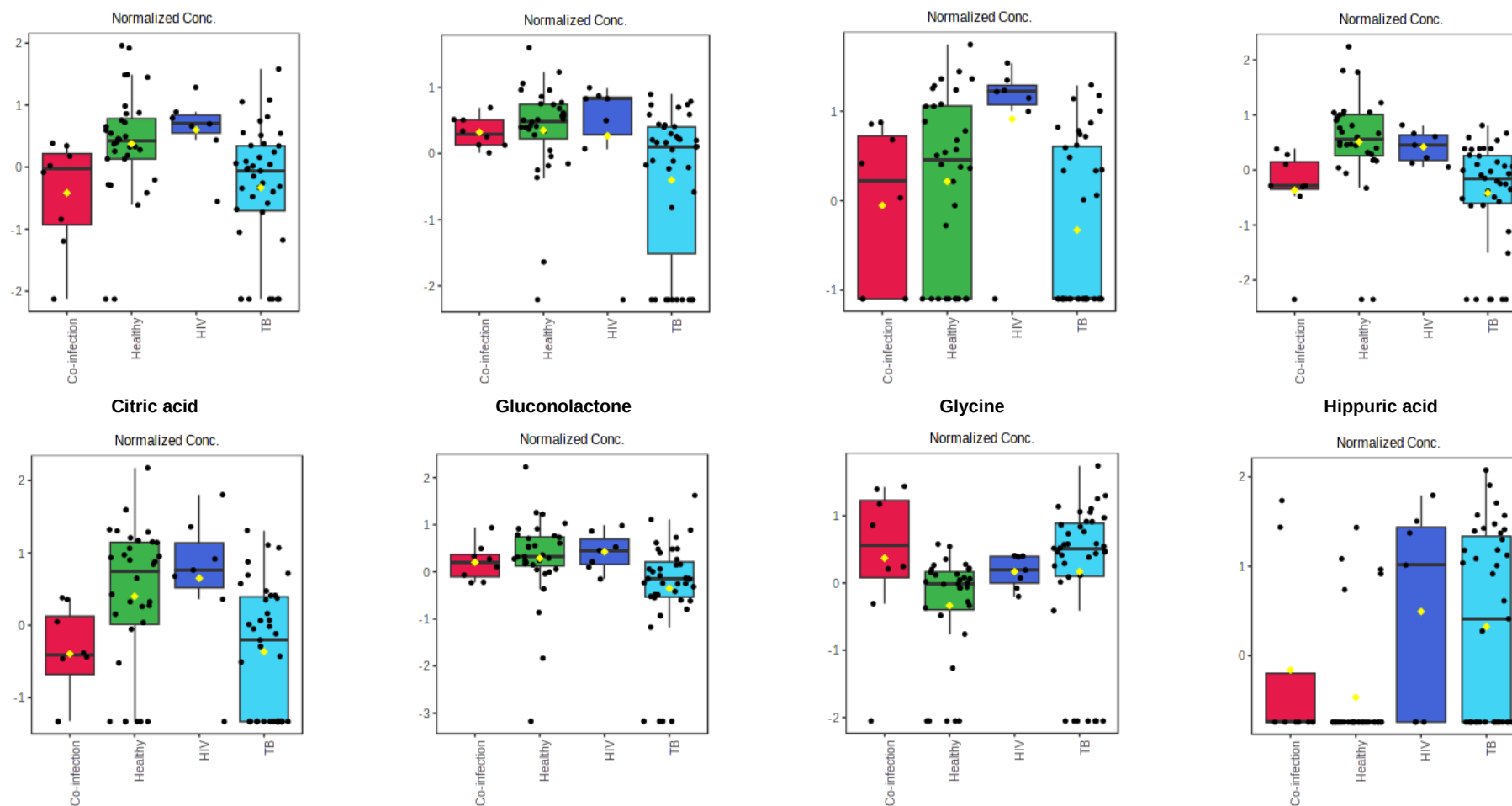


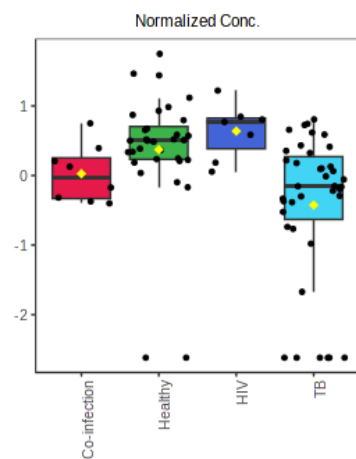
Fig. S5: AIDS-related deaths (blue line) in South Africa from 2010 to 2020, with the World Health Organization's target for 2023 represented by the orange dot. An exponential trend line was added (the dotted blue line) to depict a forward projection

APPENDIX D

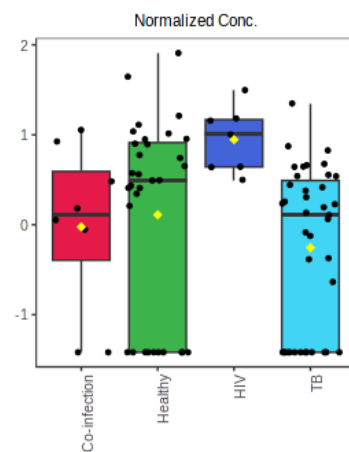
Table A1: Boxplots of significant metabolites across different groups, depicting individual samples (black dots), group averages (yellow dots), and group mean concentrations (horizontal lines within boxes).



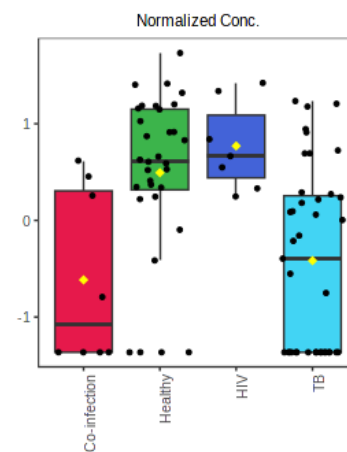
Hydracrylic acid



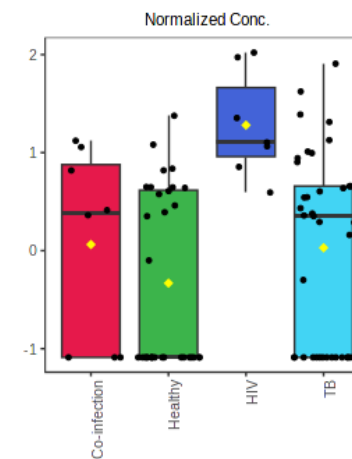
Meso-Erythritol



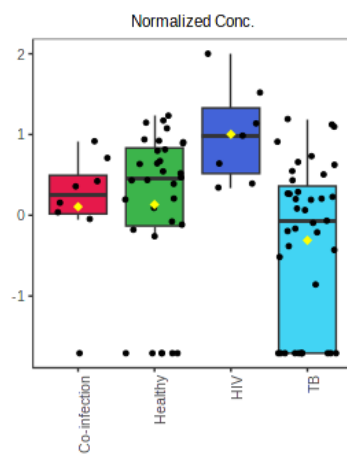
Quinolinic acid



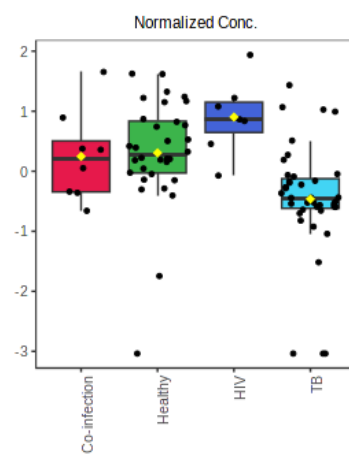
2-Methylene-butane-1,4-diol



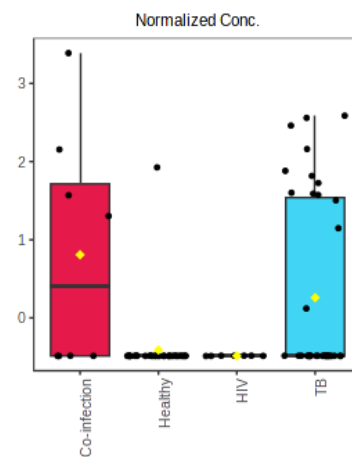
2,3-Dihydroxybutanoic-Acid



Methylmalonic acid



2,3-Butanediol



Pentanedioic acid

2-Deoxyribolactone

Propanoic acid

5-(3,4-Dimethoxy-benzyl)-pyrimidine-2,4-diamine