

A metabolomics investigation of in utero antiretroviral exposure to neonates

GP Moutloatse
20212100

Thesis submitted for the degree *Philosophiae Doctor* in
Biochemistry at the Potchefstroom Campus of the North-West
University

Promoter: Prof CJ Reinecke

October 2017

This thesis is dedicated to a phenomenal woman

Thenjiwe Beatrice Moutloatse

my Mother

“All that I am and ever hope to be is because of my dear Mother”

ACKNOWLEDGEMENTS

“I have seen something else under the sun: The race is not to the swift or the battle to the strong, nor does food come to the wise or wealth to the brilliant or favor to the learned; but time and chance happen to them all.”

Ecclesiastes 9: 11

I am continually cognisant of the fact that no man is an island. We are a sum-total of what we have learned from all who have taught us, both great and small. Although we may not be awarded the same opportunities, we are all awarded the same commodity, time. This work is a result of my academic life-time of learning and personal development. It is the collective contribution of many mentors, teachers, supporters, advisors, friends and family.

To my mother, the cradle which holds me. I am grateful for all the encouragement and wisdom that you have imparted and instilled within me. For constantly reminding me, in the words of Ralph Waldo Emerson “that envy is ignorance and imitation is suicide.” Thank you for being my beacon of light when I was caught up in my doctorate blues, and most sincerely for being my place of solace and spiritual recalibration.

To my father, a man of few words. Thank you for your unwavering love and unconditional support.

To my sister Dr. Itumeleng Moutloatse-Makine. You are and always will be a true sister’s keeper. Thank you for all your timeless efforts and constant support in my academic growth. You’re a tremendous source of motivation and encouragement but most of all a friend. To your loving Husband Dr. Brian Makine. Thank you for all the support, for being my sound board and a source of inspiration. To Ofentse and Tumisho Makine, thank you for making it easier for me to retain the kid-side of Auntie G. It’s been your innocence and child’s view of the world, along with my maturity and understanding that’s made my PhD-journey rare.

“If your actions inspire others to dream more, learn more, do more and become more, you are a leader...” JQ Adams. To Prof. Carolus Reiniecke. Thank you immensely for being a great mentor and in oftentimes for being my “Papa C”. So rare it is to meet a person, who not only is passionate about what they do but have found and love their true purpose. To me Prof you are the mirror of what hard work, and tireless dedication brings. Your love for science and metabolomics is indeed an “ART”.

To Dr. Udo Engelke and Prof. Ron Weaver. Thank you for your support, expertise and mentorship you brought into my PhD project. For making my stay at Radboud UMC an adventurous one and most of all a second home. I appreciate the lifelong NMR skills which you have imparted with me and for your instrumental contributions towards our manuscript. To Udo a special thank you for the moments you have shared with me and your family whilst in Nijmegen, I will forever cherish and hold on to the memories.

To Dr. Madeleine Bunders. Thank you for your supportive and leading role you played in obtaining this Phd. Thank you for your support at the Academic Medical Center at the University of Amsterdam, for facilitating the sample collection and patient data collection. Subsequently, to being an instrumental co-author in our manuscripts.

A special thank you go out to my colleagues and co-authors in our publications:

- Dr. Gerhard Koekemoer & Dr. Shayne Mason. Thank you for your endurance, expertise and hard-work in turning KEMREP, a 2year arduous journey into fruition. Special thanks go out even further to Dr. Shayne Mason for his willingness and unwavering support through this PhD journey. I have been lucky to have you not only as friend but to share each milestone.
- Mari van Reenen. Thank you not only for all the statistical work and bioinformatics which you put into this PhD, but for your friendship and support throughout it all
- Dr Nelus Schoeman. Thank you for the collaborative hard-work we did, and support throughout our publication
- Dr. Zander Lindeque. Thank you for your tireless support and empowering me with LC-MS-MS skills and training.

Your friendships I will take with me further into my life-long journey.

Further acknowledgements go out to Dr Graham Baker and Dr. Elisabeth Lickindorf from Kerlick Editorial Research Solutions. Thank you for your tireless efforts and support in constructing this thesis and the published manuscript into material of higher standards. Your editorial skills, critical comments and writing workshops were all highly appreciated. Not only have I skillfully been able to learn the “ART” and science of getting published, but I will forever strive to create work which is “NOUS” – New, Original, Useful & Special.

To my dearest friends, your contributions have not been taken light-heartedly. I am even more grateful for the distractions, tireless support and enthusiasm you catered for during this rigorous PhD journey. Your time and efforts have meant the world to me and for that, Thank-you.

Last but not least, special thanks go out to the academic institutions and research affiliations which made this PhD all possible:

¹Centre for Human Metabolomics, Faculty of Natural Sciences, North-West University (Potchefstroom Campus), Private Bag X6001, Potchefstroom, South Africa.

²Potchefstroom Laboratory for Inborn Errors of Metabolism Division for Biochemistry, North-West University (Potchefstroom Campus), South Africa

³Department of Experimental Medicine, Academic Medical Centre (AMC), University of Amsterdam (UvA), Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

⁴Emma Children's Hospital, AMC, University of Amsterdam. The Netherlands.

⁵Department of Pediatric Hematology, Immunology and Infectious Diseases, Emma Children's Hospital, AMC, University of Amsterdam. The Netherlands.

⁶Translational Metabolic Laboratory - 830 TML, Department Laboratory Medicine, Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands.

⁷Department of Analytical Biosciences, Leiden Academic Center for Drug Research, Leiden University, Einsteinweg 55, 2333 CC, Leiden, The Netherlands.

ABSTRACT

This thesis, with the title “*A metabolomics investigation of in utero antiretroviral exposure to neonates*”, deals with the concern over adverse metabolic consequences reported for some neonates prenatally exposed to antiretroviral treatment (HIV-ART). The pandemic due to HIV infection, the cause of AIDS, has become globally the most devastating infectious disease. Mother-to-child transmission (MTCT) of HIV during the intrapartum period became a pressing issue of the HIV pandemic. The development of ART and its use in HIV positive pregnant women to prevent mother-to-child transmission (PMTCT), is described as one of the greatest successes in medical history of the 20th century and became a standard recommendation of the WHO to combat the AIDS pandemic.

We present here the first metabolomics investigation of HIV-ART exposure to neonates, studied on cord blood taken at birth.

Chapter 1 gives the background to the study, which formed part of a larger international inter-university project, whereas Chapters 2 and 3 provide (1) a literature review of the societal, biological and biochemical aspects of HIV-ART and (2) of the use of metabolomics in the study of a biological question, respectively. The aim of the study is defined as an investigation to address the concern following treatment to PMTCT by exploring metabolic profiles in exposed neonates. The research question that was investigated was to determine whether any HIV-ART-induced metabolic perturbations are discernible during birth when the fetus is exposed to the trauma of birth hypoglycaemia.

Chapter 4 describes the use of an untargeted metabolomics [^1H nuclear magnetic resonance (NMR) spectrometry] to analyse cord blood from HIV-ART exposed and unexposed neonates, as well as serum from healthy infants. Decreased levels of 3-hydroxybutyrate and alarming increase in hypoxanthine, indicators of metabolic stress, were observed in the cord blood of exposed neonates. Although the newborns that were subject to HIV-ART exposure seem to cope with the metabolic stress of birth, the biomarkers indicated some of them to be at risk, which warrants further monitoring.

Chapter 5 focuses on the analytical aspects of liquid chromatography–mass spectrometry (LC-MS) required for a method of higher sensitivity to complement the NMR investigation. Triple quad mass spectrometry (LC-QQQ) proved to be a method of choice given its high repeatability for a targeted analysis of acylcarnitines and amino acids in cord blood.

Chapter 6 provides results that support the benefits of ART for HIV-infected pregnant women as recommended by the WHO, but the amino acid profiles of cord blood, and to a lesser extent of the acylcarnitines, add to the observed metabolic risks with potential adverse consequences for

some infants exposed *in utero* to HIV-ART, which might affect their future health. Results from a lipidomics study of an independent cohort of HIV-ART exposed newborns (a complementary part of the inter-university study) furthermore support the conclusions from the present study.

Chapter 7 discusses the findings of the investigation presented in this thesis, indicates some of its limitations and concludes with a mini-review on the potential of cord blood for metabolomics studies in the field of neonatology.

Key words: HIV-antiretroviral therapy (HIV-ART); Prevention of mother-to-child transmission (PMTCT); Cord blood metabolome; Transient hyperglycaemia; Metabolic risk; ¹H nuclear magnetic resonance spectroscopy; Liquid chromatography–mass spectrometry.

Publications:

The following papers have been published in peer-reviewed journals or are under review, based on results presented in this thesis* or from a collaborative investigation#:

#Mason, S., Moutloatse, G.P., Marceline van Furth, A., Solomons, R., van Reenen, M., Reinecke, C., & Koekemoer, G. (2014). KEMREP: A new qualitative method for the assessment of an analyst's ability to generate a metabolomics data matrix by gas chromatography–mass spectrometry. *Current Metabolomics*, 2(1), 15–26.

*Moutloatse, G.P., Bunders, M.J., van Reenen, M., Mason, S., Kuijpers, T.W., Engelke, U. F.H. & Reinecke, C.J. (2016). Metabolic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood. *Metabolomics*, 12(175), 1–14.

*Moutloatse, G.P., Schoeman, J.C., Lindeque, Z., Van Reenen, M., Hankemeier, T., Bunders, M.J., & Reinecke, C.J. (2017). Metabolic risks of neonates at birth following in utero exposure to HIV-ART: the amino acid profile of cord blood. *Metabolomics*, 13(89), 1–11.

#Schoeman, J. C., Moutloatse, G. P., Harms, A. C., Vreeken, R. J., Scherpbier, H. J., Van Leeuwen, L., et al. (2017). Fetal metabolic stress disrupts immune homeostasis and induces pro-inflammatory responses in HIV-1 and cART-exposed infants. *The Journal of Infectious Diseases*.

*Moutloatse, G.P. and Reinecke, C.J. ¹H-NMR metabolomics studies on neonatal cord blood – Submitted for publication to *Current Metabolomics*.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	i
ABSTRACT.....	ERROR! BOOKMARK NOT DEFINED.
CHAPTER 1 BACKGROUND.....	1
REFERENCES.....	6
CHAPTER 2 IN UTERO EXPOSURE TO HIV-ART – A LITERATURE REVIEW.....	8
2.1 Origins and History of HIV/AIDS: A brief overview	8
2.2 Epidemiology and South African’s brief history	10
2.2.1 HIV in the South African socio-political context.....	12
2.2.2 Denialism in a post-apartheid health system	13
2.2.3 The mother-to-child transmission of HIV	13
2.3 Mother-to-child transmission (MTCT) and risk factors.....	15
2.3.1 Timing of MTCT HIV transmission	15
2.3.2 Maternal factors	16
2.3.3 Obstetrical factors	17
2.4 Prevention of mother-to-child transmission (PMTCT).....	18
2.4.1 Setting the stage	18
2.4.2 WHO PMTCT recommendations and strategy for pregnant women.....	20
2.5 Long-term consequences of HIV-ART exposure	21
2.6 Metabolic studies of HIV and ART exposure.....	22
2.6.1 Metabolic and mitochondrial indicators of HIV and ART exposure	22

2.6.2	Metabolomic studies of HIV	30
2.7	Aim and research question for the investigation.....	36
REFERENCES.....		37
CHAPTER 3 METABOLOMICS METHODOLOGY		49
3.1	Outline	49
3.2	Metabolomics: from small molecules to data to big ideas	49
3.3	Approaches to metabolomics	53
3.3.1	Background	53
3.4	Approaches to metabolomics	54
3.4.1	Define: Research question and design	54
3.4.2	Acquire: Sample processing and measurement	62
3.4.3	Discover: Data pre-treatment metabolite annotation.....	65
3.4.4	Interpretation: From information to hypothesis formulation to betterment.....	68
3.4.5	Disseminate: New/translational knowledge.....	71
REFERENCES.....		73
CHAPTER 4 METABOLIC RISKS AT BIRTH FOLLOWING HIV-ART EXPOSURE: INSIGHTS FROM AN UNTARGETED METABOLOMICS STUDY		80
4.1	Introduction	80
4.1.1	Birth asphyxia and extra-uterine adaptation.....	80
4.1.2	Metabolic perturbations at birth and extra-uterine adaptation	81
4.1.3	The untargeted metabolomics study of cord blood from HIV-ART exposed and unexposed neonates.....	82
4.2	Introduction	86

4.3	Methods	89
4.3.1	Study design	89
4.3.2	Samples and the study groups.....	91
4.3.3	Sample preparation and metabolomics analysis.....	93
4.4	Statistical analysis	94
4.5	Results and Discussion.....	95
4.5.1	Data generation and case selection	95
4.5.2	Metabolite profile of cord blood of neonates unexposed to HIV-ART at birth	96
4.5.3	Metabolite profile of cord blood of neonates exposed to HIV-ART at birth	99
4.5.3.1	Metabolic risks seen at birth for neonates exposed to HIV-ART in utero.....	103
4.5.3.2	Significant discriminating metabolites on in utero HIV exposure	106
4.6	Concluding remarks	107
REFERENCES.....		110
 CHAPTER 5 REPEATABILITY IN THE GENERATION OF A MASS-SPECTROMETRIC-BASED METABOLOMICS DATA MATRIX		
5.1 Extended metabolite identification in HIV-ART cord blood of neonates ...		117
5.2	Metabolomics data reliability	119
5.2.1	Generation of a metabolomics matrix	120
5.2.2	The analyst.....	121
5.2.2.1	Sample selection and preparation.....	121
5.2.2.2	Analytical instrumentation: Mass spectrometry.....	122
5.2.2.3	Quality control	122
5.3	KEMREP: A new method for the assessment of an analyst's ability	124

5.3.1	Background	124
5.3.2	Experimental method and results.....	126
5.3.3	The KEMREP publication.....	127
5.4	Application of KEMREP in LC-QQQ	128
5.4.1	Experimenatal approach	128
5.4.2	Reagents.....	129
5.4.3	Sample preparation.....	129
5.4.4	LC-QQQ acylcarnitine and amino acid analysis.....	130
5.4.5	Data extraction and normalisation of acylcarnitines and amino acids.....	130
5.4.6	Statistical analysis.....	131
5.4.7	Conclusions.....	133
	REFERENCES.....	135

CHAPTER 6 METABOLIC RISKS AT BIRTH FOLLOWING HIV-ART EXPOSURE: INSIGHTS FROM SEMI-TARGETED METABOLOMICS STUDIES 137

6.1	A semi-targeted metabolomics analysis of cord blood acylcarnitines.....	137
6.1.1	Introduction	137
6.1.2	Materials and methods.....	139
6.1.2.1	Reagents.....	139
6.1.2.2	The experimental subjects for the metabolomics analysis.....	139
6.1.2.3	Analysis of acylcarnitines using liquid chromatography–tandem mass spectrometry	139
6.1.2.4	Data processing and statistical analysis	140
6.1.3	Results and discussion	140

6.2	A semi-targeted metabolomics analysis of cord blood amino acids	143
6.2.1	Background	143
6.2.2	The amino acid profile of cord blood from HIV-ART exposed and unexposed neonates	144
6.2.2.1	Introduction	146
6.2.2.2	Materials and methods.....	148
6.2.2.2.1	Reagents and stable isotope standard.....	148
6.2.2.3	Cases, Sample Collection and Preparation	148
6.2.2.4	Experimental procedures	149
6.2.2.4.1	<i>Sample extractions</i>	149
6.2.2.4.2	<i>LC-QQQ analysis for cohort-1</i>	149
6.2.2.4.3	<i>UPLS-MS/MS analysis for cohort-2</i>	150
6.2.2.4.4	<i>Statistical analysis</i>	151
6.2.2.5	Results	151
6.2.2.6	Discussion.....	158
6.2.2.7	Concluding views	160
6.3	A semi-targeted metabolomics analysis of cord blood lipids	162
6.3.1	Background	162
6.3.2	Lipidomics and pro-inflammatory immune responses in cART-exposed neonates	162
6.4	Conclusions.....	163
REFERENCES.....		165
CHAPTER 7 DISCUSSION.....		171

7.1	The aim of the study	171
7.1.1	The appropriateness of metabolomics for analysis of cord blood	171
7.1.2	Answering the research question	172
7.1.3	Views on the aim of the study, and of the contribution it makes	173
7.1.4	The limitations of my study.....	174
7.2	The aim defined by the international strategic planning group	175
7.3	Future prospects	178
7.3.1	Introduction	179
7.3.2	The cord blood metabolome	179
7.3.3	Metabolomics insights on pregnancy complications	182
7.3.3.1	Intra uterine growth restriction.....	182
7.3.3.2	Preterm infants.....	184
7.3.3.3	Low birth weight neonates	186
7.3.3.4	Neonatal asphyxia.....	187
7.3.4	Fetal-maternal disease conditions	189
7.3.4.1	Gestational diabetes mellitus	189
7.3.4.2	HIV-ART exposure	191
7.3.5	Future prospects	193
REFERENCES.....		196
ANNEXURE 1:.....		206
REFERENCES.....		233
ANNEXURE 2: KEMREP.....		235
ANNEXURE 3:.....		243
REFERENCES.....		255

ANNEXURE 4: COPYRIGHT LICENCING AGREEMENTS	256
---	------------

LIST OF TABLES

Table 2-1:	Estimation of timing of MTCT in non-breastfeeding and breastfeeding populations (adapted from Kourtis <i>et al.</i> 2006)	16
Table 2-2:	Antiretroviral drugs used on pregnant HIV-infected women (Adapted from Simon <i>et al.</i> 2006)	19
Table 4-1:	Clinical data details of the neonates <i>in utero</i> exposed to HIV-AR.....	93
Table 4-2:	Quantitative data on the HIV-ART-exposed and unexposed neonates for important metabolites identified from the spectra corresponding to the two experimental groups. The chemical shift for the specific bins used for the quantification is shown in brackets after the name of the metabolite	102
Table 6-1:	Acylcarnitines identified from the two experimental groups analyzed using multiple reaction monitoring (MRM).....	142
Table 6-2:	Summary of the mean concentrations and standard deviations of the amino acids with univariate and multivariate parameters for cohort-1.....	155
Table 6-3:	Summary of the relative concentrations and standard deviations of the amino acids with univariate and multivariate parameters for cohort-2.....	156

Table 7-1:	A selection of metabolites from cord blood detected by ^1H NMR-based metabolomics	
		181

LIST OF FIGURES

Figure 1-1:	Flow diagram of the experimental design developed in 2012 for the international inter-university research project on the metabolic consequences of HIV-ART exposure of neonates <i>in utero</i> , as studied through metabolomics investigations of cord blood.....	4
Figure 2-1:	Phylogenetic relationships among representative primate lentiviruses. For HIV-1 the group of origin (red dashed box), and for HIV-2 the subtype (blue dashed box), are indicated both on the left and right section of the figure. SIV strains have their host species of origin: cpz, chimpanzee (HIV-1) and mac, macaque; sm, sooty mangabey (HIV-2). (Adapted from Sharp <i>et al.</i> 2001; Williams 2012)	10
Figure 2-2:	Adults and children estimated to be living with HIV in 2015 by WHO region (WHO 2016)	11
Figure 2-3:	A conceptual representation of the effect of HIV and ART on mitochondrial metabolism. Antiretroviral (ARV) exposure is shown to affect lipid (dyslipidemia) and carbohydrate (lactic acidosis) metabolism, fatty acid oxidation (FAO) and lead to the formation of acetyl-CoA, and mitochondrial (oxidative phosphorylation) metabolism. HIV infection itself is primarily associated with a variety of effects on mitochondrial DNA (mtDNA), with increased generation of reactive oxygen species (ROS), dysfunctional lipid metabolism and some minor effects on phenylalanine (PHE) as well as metabolism of some other amino acids. (This diagram is an elaboration of Figure 1 with permission from Kirmse <i>et al.</i> 2013b)	24
Figure 3-1:	Number of manuscripts returned from a PubMed search using the terms 'Metabolomics' and 'Metabonomics' [Adapted from Powers, (2002- 2015)	

and Kell & Oliver 2014-2015)]	51
-------------------------------------	----

Figure 3-2:	A conceptual representation of the evolution of the metabolic phenotype as a consequence of an <i>in utero</i> exposure (or no exposure) to maternal HIV-ART. Environment 1 relates to the genetic and epigenetic characteristics of the father and mother, transferred at conception to the fertilized ovum. Environment 2 indicates the prenatal environment of the fetus without exposure (controls) or exposed to HIV-positive mothers treated with ART. Environment 3 indicates the potential long-term consequences of HIV-ART on the infants, as they develop towards adulthood.....	52
-------------	---	----

Figure 3-3:	Schematic representation of the workflow for the metabolomics investigation. The design is presented as an overview of a pipeline of a metabolomics platform ranging from the design, performance, analysis and interpretational aspects, dissemination of metabolomics experiments (highlighted by rectangular boxes shown to the left of the figure) and their attendant requirements and multidisciplinary involvements (highlighted by the rectangular boxes shown to the right of the figure)..	55
-------------	--	----

Figure 3-4:	Scientific advance may be seen as an interactive cycle linking knowledge and observations. (Reproduced with permission from Kell & Oliver 2004).	69
-------------	--	----

Figure 4-1:	A schematic representation of the investigation workflow to generate data, to identify a metabolite profile of neonatal hypoglycaemia and to summarize the assessment of the effect of HIV-ART exposure on neonates at delivery. Circled numbers 1–7 indicate the sequence of methodological stages followed in this investigation: stage 1 represents the process towards data generation; stages 2, 3 and 7 were aimed at revealing the characteristic global metabolite profile of transient neonatal hypoglycaemia; stages 4–7 focused on identifying metabolic risks, if any,
-------------	--

for neonates *in utero* exposed to HIV-ART. An asterisk (*) in the flow diagram indicates main conclusions formulated for the respective stages)

.....90

Figure 4-2: Statistical analyses to identify metabolites being important during transitional neonatal hypoglycemia in neonates exposed or not exposed to HIV-ART *in utero*. (a) PCA score plots of infant and unexposed neonatal groups based on the normalized data from the 113 bins. (b) PLS-DA score plots of infant and unexposed neonatal groups based on the 113 bins. (c) PLS-DA score plots of infant and unexposed newborns based on the normalized data from the 53 bins. (d) PCA score plots of unexposed and exposed newborns based on the 113 bins. (e) PLS-DA score plots of exposed and unexposed neonatal groups based on the 113 bins. (f): PLS-DA score plots of unexposed and exposed newborns based on the normalized data from the 53 bins (g) PCA S-plot for the infant and unexposed neonatal groups, using spectral data from the 113 bins. (h) Volcano plot of the 15 quantified metabolites from the neonatal groups exposed or unexposed to HIV-ART diagram indicates main conclusions formulated for the respective stages)..... 97

Figure 4-3: Graphs showing important metabolites indicating transient neonatal hypoglycaemia. Indicated in the figure are: Infants relative to unexposed neonates (3a–3c) and HIV-ART exposure cases relative to infants or unexposed cases (3d–3e). Values for all individual cases are shown as dots, while the squared area represents the 95 % confidence interval (red) and 1 standard deviation (blue) of the mean (red line). The statistical significance of differences (i.e. *p*-values) is indicated for every group relative to the other, in the upper section of each set of graphs (*: *p* < 0.05; **: *p* < 0.01; ***: *p* < 0.001). The threshold of fasting hypoglycaemia (2.5 mM) is indicated by the red dotted line in (b). The profiles of four newborns with extreme hypoglycemia are highlighted, shown in orange and numbered 1 and 2 for the two unexposed and 3 and 4 for the two HIV-ART-exposed cases. The values for these four cases are likewise

indicated for the remaining metabolites (3d and 3e), as well as in Figure 4-4 100

Figure 4-4: Regression lines of potential ATP generation and OR plots of important informative metabolites for all individual case studies. (a) Data on low (0–2.5 mM) cord blood glucose for unexposed neonates according to $[ATP]_{LA/3HB}$ (lower curve – blue) and $[ATP]_{GI/LA/3HB}$ (upper curve – green). (b) Individual values for the HIV-ART-exposed neonates (red) to compare ATP generation according to $[ATP]_{LA/3HB}$ against low cord blood glucose for the unexposed neonates (solid line – blue) including its 90 % confidence intervals (dotted lines – blue). (c) Odds ratio plots with 95 % confidence intervals for 3-hydroxybutyric acid. (d) Odds ratio plots with 95 % confidence intervals for hypoxanthine. The estimated OR is indicated by the red diamond on the horizontal line and the vertical dashed line indicates significance if CI does not contain 1.0..... 105

Figure 5-1: Data processing and the creation of a metabolomics data matrix (adapted from Jonsson *et al.* 2005) 120

Figure 5-2: Flow diagram illustrating the experimental procedures followed by analysts in the metabolomics pipeline and showing the role subsequently played by KEMREP 125

Figure 5-3: Overlays of density plots of 5 repeat aliquots from a single urine sample. (A) First attempt to generate the GS-MS data. (B) Second attempt to generate the GS-MS data by the same analyst. All analyses were conducted using the identical reagents, standard operating procedure (SOP) and the same GC-MS

	apparatus.....	
		126
Figure 5-4:	Repeatability results: CV distribution, density methods and kernel density estimates of an LC-QQQ acylcarnitine and amino acid analysis done on two days. Day 1 repeatability analysis is represented in 5-4a and 5-4b. Following Day 3 analysis are presented in 5-4c and 5-4d.....	132
Figure 6-1:	Statistical analysis of carnitines of neonates exposed (red dots) and unexposed (blue dots) to HIV-ART <i>in utero</i> as detected with untargeted LC-MS analyses with the red and blue triangles representing the center of the PCA. a: PCA score plot of normalized data of <i>in utero</i> HIV-ART-exposed infants against their unexposed controls. b: PLS-DA score plots of normalized data of <i>in utero</i> HIV-ART-exposed infants against unexposed controls	141
Figure 6-2:	Multivariate statistical analyses to identify metabolites important in neonates exposed or not exposed to HIV-ART <i>in utero</i> . (a) PCA score plots of exposed and unexposed neonatal groups based on the quantified (μ moles amino acid) normalized amino acid data from the present experimental groups. (b) PCA score plots of the comparative exposed and unexposed neonatal groups based on semi-quantitative normalized amino acid data. (c) PLS-DA score plots of the same data as shown in a. (d) PLS-DA score plots of the same data as shown in b. Data from all HIV-exposed neonates are shown in red and for the unexposed cases in blue. The respective centroids of the 90% confidence ellipsoids are shown as red and blue triangles.....	153
Figure 6-3:	Statistical analysis of carnitines of neonates exposed (red dots) and unexposed (blue dots) to HIV-ART <i>in utero</i> as detected with untargeted	

LC-MS analyses with the red and blue triangles representing the center of the PCA. a: PCA score plot of normalized data of *in utero* HIV-ART-exposed infants against their unexposed controls. b: PLS-DA score plots of normalized data of *in utero* HIV-ART-exposed infants against unexposed controls.....

154

Figure 6-4: Graphs showing aspartic acid profiles from Cohort 1 (quantified concentrations) and cohort-2 (relative concentrations). Indicated in the figure are: Two infants (indicated by red squares) of HIV-ART exposure from cohort-1 (a) and cohort-2 (b) showing very high cord blood aspartic levels (~3.5 times increased) relative to the mean values of the corresponding exposed neonates. Values for all other individual cases are shown as black dots. The boxed areas represent the 95% confidence interval (orange) and 1 standard deviation (blue) about the mean (black line).

157

Figure 6-5: Graphs showing the quantified concentrations for alanine (a), glycine (b) and tyrosine (c) as observed in the controls and in cohort-1. A neonate (indicated by a red dot) of the HIV-ART exposure group appears to be at risk for hypoglycaemic allostasis, based on clinical parameters. Values for all other individual cases are shown as black dots. The boxed areas represent the 95% confidence interval (orange) and 1 standard deviation (blue) about the mean (black line).....

158

Figure 7-1: A proposed representation of perturbations of homeostasis at and following birth (Original model derived from Moutloatse *et al*, 2014 – poster presentation).....

176

CHAPTER 1 BACKGROUND

The pandemic of HIV infection, the cause of the acquired immunodeficiency syndrome (AIDS), is globally a most devastating disease, responsible for a cumulative of nearly over 75 million infections since the world became aware of it (Fauci 1999; UNAIDS 2002). AIDS was first noticed and reported in the United States (U.S.) in the period from October 1980 to May 1981, when actively practising homosexual men were presenting with Kaposi's sarcoma and opportunistic infections such as *Pneumocystis carinii* pneumonia (Friedman-Kien *et al.* 1981; CDC 1981). Since then, HIV/AIDS has exploded in the human population in successive waves in various regions in the world (Fauci 1999). Mother-to-child transmission (MTCT) of human immunodeficiency virus type 1 (HIV-1)¹, largely transmitted during the intrapartum period (Read *et al.* 1999), became a pressing issue of the HIV pandemic. The development of antiretroviral therapy (ART), and its use in pregnant women with HIV infection to prevent mother-child transmission (PMTCT), is described as nothing short of one of the most extraordinary successes in the medical history of the 20th century (Fauci *et al.* 2013).

In 2011 the Joint United Nations Programme on HIV and AIDS (UNAIDS), and other global policy makers such as, the U.S. President's Emergency Plan for AIDS Relief (PEPFAR), set forth a global strategy to eliminate MTCT and keep HIV-positive mothers alive. This operation mainly focused on the expansion of PMTCT especially in low-income, HIV-prevalent countries such as those of sub-Saharan Africa (UNAIDS 2011). In the 2015 amended guidelines for PMTCT, the World Health Organization (WHO) recommends that countries adopt a lifelong, universal ART antenatal treatment plan for all pregnant and breastfeeding women (WHO 2015). This lifelong plan will improve the mother's health outcomes during her pregnancy, reducing mortality while also, remarkably, lowering the chances of vertical MTCT of HIV. However, offering lifelong ART treatment to all pregnant and breastfeeding woman will increase the duration of intrauterine ART exposure to the fetus, leading to cumulative drug toxicity, which can imply increased risks of adverse health outcomes in adulthood (Ahmed *et al.* 2013; Jao & Abrams 2014).

¹ All variants of HIV (e.g. HIV-1) hereafter referred to HIV unless otherwise required for specific purposed, as in Chapter 2.

In South Africa, 30% of uninfected children are born to HIV-positive mothers. There is therefore an increasing prevalence of uninfected, yet HIV-exposed children (Slogrove *et al.* 2016). Although the use of antenatal ART has been successful in restraining the vertical–perinatal transmission of HIV, the long-term effects of its use during pregnancy are unknown. In addition, there is a growing concern that HIV-ART may affect metabolomic functioning and impair fetal and infant development (Barret *et al.* 2003). A wide range of literature has presented evidence of *in utero* exposure to nucleoside reverse transcriptase inhibitors (NRTIs), a class of antiretroviral drugs, linked to the occurrences of mitochondrial dysfunction, abnormal mitochondrial DNA, altered oxidative phosphorylation as well as morphologically compromised mitochondria in HIV-uninfected children born to HIV-infected women (Poirier *et al.* 2015; Brogly *et al.* 2007; Blanche *et al.* 1999).

Furthermore, reported clinical manifestations of infants exposed to NRTIs *in utero* include metabolic disruption, cardiac and neurological malformation as well as transient lactic acidemia (Poirier *et al.* 2015; Alimenti *et al.* 2003). The continued intrauterine exposure of the fetus to combination ART (cART) during its vulnerable and most critical period of development (fetal programming), warrants close monitoring (Jao & Abrams 2014). In this regard, Dr M. Bunders, paediatrician at the Department of Experimental Immunology, Academic Medical Centre (AMC) in Amsterdam, forewarned:

“The benefit of ART in the prevention of MTCT is beyond doubt, but there are reports on adverse effects on pregnancy outcome and infant currently also from high prevalence regions. Further research regarding safety is urgently required, as the number of pregnant women on ART and exposed uninfected infants is rapidly increasing.” (Newell & Bunders 2013)

Dr Bunders has an established research collaboration with paediatricians at the University of Cape Town. Since the 2010s, she became progressively aware of the importance of metabolomics in the study of affected human conditions, and knew of the launching of metabolomics platforms in several countries, including South Africa and the Netherlands. Dr Bunders was introduced to my promoter, Professor Reinecke, by Dr Graham Baker, former Editor of the *South African Journal of Science* and acquainted with Dr Bunders on a personal level. Following an informal meeting between the two scientists, Dr Bunders took the initiative to organize a

meeting convened at the Department of Analytical Biosciences, University of Leiden, the site of a newly established metabolomics platform in the Netherlands. Amongst others, Dr Bunders, Professor Reinecke (Manager of the South African Metabolomics Platform, hosted at the Potchefstroom Campus of North-West University) and Professor T. Hankemeier [(Director of the BioMedical Metabolomics Facility Leiden (BMFL), hosted by the University of Leiden)] embarked on a strategic planning exercise to investigate the concerns over *in utero* HIV-ART exposure through a comprehensive metabolomics approach. As the two platforms involved did not have an NMR facility – which is a key instrument for untargeted metabolomics studies – it was agreed to include Professor R.A. Wevers, Head of the NMR facility for Clinical Chemistry at the Nijmegen University Medical Centre (NUMC) in the programme, given the cooperation that already existed between the research groups of Professors Reinecke and Wevers.

A comprehensive experimental design for a metabolomics investigation on HIV-ART exposure – the first of its kind internationally – was developed at the strategic planning session, shown in Figure 1-1. The experimental approach included the three major analytical platforms (NMR, GC-MS and LC-MS). Clinical description and samples for the study would be provided by Dr Bunders, previously collected at the AMC for a retrospective study on immunological characteristics of neonates exposed to HIV-ART. It was also agreed that funding for the project would be generated from the Technology Innovation Agency (TIA) in South Africa for the South African–Nijmegen component and from the Netherlands Metabolomics Consortium for the Dutch (Amsterdam–Leiden) component. Close cooperation between the project leaders of independent research endeavours was agreed upon by the group, as each party already had an extensive research agenda on other, unrelated projects. Mr J.C. (Nelus) Schoeman, an alumnus from NWU and full-time PhD student at Leiden University under the supervision of Professor Hankemeier, and myself under the supervision of Professor Reinecke for a PhD at the North-West University (NWU), would participate as the primary researchers in the project. It was my obligation to prepare all cord blood samples for the three platforms and to generate the clinical information on the mothers and neonates from the records at the AMC, conducted under the supervision of Dr Bunders.

The bioinformatics co-operation – from the planning to the data analysis phases – would be done for the projects at the two universities involved.

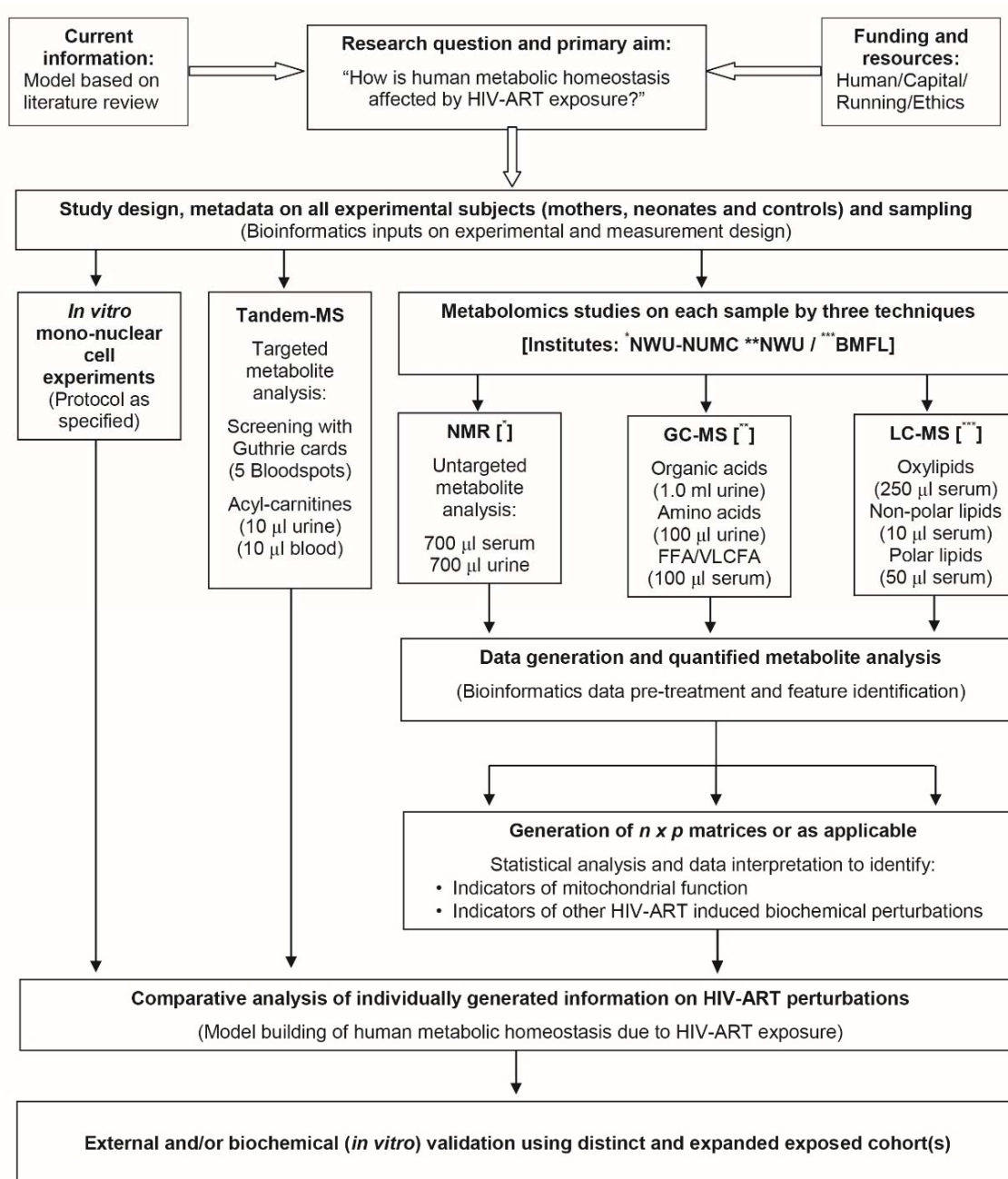


Figure 1–1: Flow diagram of the experimental design developed in 2012 for the international inter-university research project on the metabolic consequences of HIV-ART exposure of neonates *in utero*, as studied through metabolomics investigations of cord blood.

Ethical approval for the study was provided by the Ethical Committee of the Medical Faculty of the University of Amsterdam; Dr Bunders and Professor Hankemeier gave written recommendations to TIA and the NWU for the combined investigation.

In 2013, I motivated to the Faculty Board of the Faculty of Natural Sciences at NWU for my enrolment for a PhD study in Biochemistry, which was approved for a doctoral thesis under the title "*A metabolomics investigation of in utero antiretroviral exposure to neonates*". In this thesis, I give a conceptual representation of the presumed underlying effect of HIV and ART on mitochondrial function (Chapter 2, Figure 2-3), the main target for ART-induced dysfunction as it emerged from my literature study. The literature review culminated in a concise aim of the investigation envisaged for the thesis (Chapter 2, section 2.7). Given the complex experimental protocol, I did not develop a common model for the measurement design of the study, but rather devised a schematic work-flow, which gives an overview of a pipeline used in metabolomics platforms in general. The work-flow covers the full range of a metabolomics investigation, from the phase of initial design towards performance, analysis and interpretational aspects and finally to modes for dissemination of findings from the investigation (Chapter 3, Figure 3-3).

To conclude this background: Mr Schoeman completed his contribution to the project through a thesis entitled "*Virus-host metabolic interactions: using metabolomics to probe oxidative stress, inflammation and systemic immunity*". He successfully defended his thesis in public and the PhD degree of the University of Leiden was awarded to him on 20 December 2016. The main outcome of his investigation on the effect of *in utero* HIV-ART exposure is presented in Chapter 7 of his thesis and submitted as a manuscript for publication to *Nature Communications*, which is presently under review (Abstract included in Chapter 6 of my thesis). Contributions from my part of the project appeared in print in the November 2016 edition of *Metabolomics* (Chapter 4, section 4.2), as well as two further manuscripts submitted for publication to *Metabolomics* and *Current Metabolomics*, covered in the thesis in Chapters 4 and 7, respectively. Our main conclusions on the effect of HIV-ART exposure are discussed in Chapter 7.

REFERENCES

- Ahmed, S., Kim, M. H., & Abrams, E. J. (2013). Risks and benefits of lifelong antiretroviral treatment for pregnant and breastfeeding women: a review of the evidence for the Option B+ approach. *Current Opinion in HIV and AIDS*, 8(5), 474–489.
- Alimenti, A., Burdge, D. R., Ogilvie, G. S., Money, D. M., & Forbes, J. C. (2003). Lactic acidemia in human immunodeficiency virus-uninfected infants exposed to perinatal antiretroviral therapy. *The Pediatric Infectious Disease Journal*, 22(9), 782–788.
- Barret, B., Tardieu, M., Rustin, P., Lacroix, C., Chabrol, B., Desguerre, I., et al. & French Perinatal Cohort Study Group. (2003). Persistent mitochondrial dysfunction in HIV-1-exposed but uninfected infants: clinical screening in a large prospective cohort. *AIDS*, 17(12), 1769–1785.
- Blanche, S., Tardieu, M., Rustin, P., Slama, A., Barret, B., Firtion, G., et al. (1999). Persistent mitochondrial dysfunction and perinatal exposure to antiretroviral nucleoside analogues. *Lancet*, 354, 1084–1089.
- Brogly, S. B., Ylitalo, N., Mofenson, L. M., Oleske, J., Van Dyke, R., Crain, M. J., et al. (2007). In utero nucleoside reverse transcriptase inhibitor exposure and signs of possible mitochondrial dysfunction in HIV-uninfected children. *AIDS*, 21(8), 929–938.
- CDC. 1981. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men—New York City and California. *MMWR Morbidity and Mortality Weekly Report*, 30(25), 305–308.
- Fauci, A. S. (1999). The AIDS epidemic—considerations for the 21st century. *New England Journal of Medicine*, 341(14), 1046–1050.
- Fauci, A. S., Folkers, G. K., & Dieffenbach, C. W. (2013). HIV-AIDS: much accomplished, much to do. *Nature Immunology*, 14(11), 1104–1107.
- Friedman-Kien, A. E., Laubenstein, L., Marmor, M., Hymes, K., Green, J., Ragaz, A., et al. (1981). Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men – New York City and California. *MMWR. Morbidity and Mortality Weekly Report*, 30(25), 305–308.

Jao, J., & Abrams, E. J. (2014). Metabolic complications of in utero maternal HIV and antiretroviral exposure in HIV-exposed infants. *The Pediatric Infectious Disease Journal*, 33(7), 734–740.

Newell, M. L., & Bunders, M. J. (2013). Safety of antiretroviral drugs in pregnancy and breastfeeding for mother and child. *Current Opinion in HIV and AIDS*, 8(5), 504–510.

Poirier, M.C., Gibbons, A. T., Rugeles, M. T., Andere-Schmutz, I., & Blanche, S. (2015). Fetal consequences of maternal antiretroviral nucleoside reverse transcriptase inhibitor use in human and nonhuman primate pregnancy. *Current Opinion in Pediatrics*, 27(2), 233–239.

Read, J. S., Scherpbier, H. J., & Boer, K. (1999). The mode of delivery and the risk of vertical transmission of human immunodeficiency virus type 1: a meta-analysis of 15 prospective cohort studies: The International Perinatal HIV Group. *The New England Journal of Medicine*, 340, 977–987.

Slogrove, A. L., Goetghebuer, T., Cotton, M. F., Singe, J., & Bettinger, J. A. (2016). Pattern of infectious morbidity in HIV-exposed uninfected infants and children. *Frontiers in Immunology*, 7(164), 1–8.

UNAIDS. (2002). AIDS epidemic update. URL: http://data.unaids.org/publications/IRC-pub03/epiupdate2002_en.pdf. (Accessed September 2016).

UNAIDS. (2011). Countdown to zero: Global plan towards the elimination of new HIV infections among children by 2015 and keeping their mothers alive. URL: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/20110609_jc2137_global-plan-elimination-hiv-children_en.pdf. (Accessed September 2016).

WHO. (2015). ANTIRETROVIRAL, T. S. (2015). Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. (Available: <http://www.emtct-iatt.org/wp-content/uploads/2015/09/WHO-Guidelines-on-When-to-Start-ART-and-PrEP-September-2015.pdf>)

CHAPTER 2 IN UTERO EXPOSURE TO HIV-ART – A LITERATURE REVIEW

2.1 Origins and History of HIV/AIDS: A brief overview

HIV/AIDS, one of the most destructive global pandemics, in terms of human mortality, has undoubtedly become the defining medical and public health threat of our generation and ranks globally among the most devastating infectious diseases in history (Fauci 2003). A retrovirus, now notorious as human immunodeficiency virus type 1 (HIV-1), classified in the genus *Lentivirus* of the family *Retroviridae*, is the major cause of the corresponding acquired immunodeficiency syndrome (AIDS) (Barré-Sinoussi *et al.* 1983; Popovic *et al.* 1984).

The outbreak of AIDS was officially and formally recognized in patients in the USA in 1981 (CDC 1981). Groups at risk included homosexual men, injection-drug users, haemophiliacs who underwent blood transfusions as well as immigrants to the country who originated from Haiti and Central Africa. The rapid spread of the disease was documented in individuals with no known cause of impaired immunity and showed symptoms of rare opportunistic infections that were known to occur in people with highly compromised immune systems and succumbed to rare malignancies (Popovic *et al.* 1984). In the following years 1983–84 a retrovirus named HIV, initially known as lymphadenopathy-associated virus (LAV) and the human T-cell lymphotropic virus type III (HTLV-III), respectively, was deemed to be the cause (Hammett *et al.* 2016). It was, however, the complete sequencing of the LAV 9193-nucleotide by Wain-Hobson *et al.* (1985) that led them to question further the origin of the virus. In fact, in their description of this result Wain-Hobson and colleagues pointed out that the genetic characteristics of LAV were similar to other retroviruses and considered that HIV could be from a group of retroviruses from which it evolved. In their paper in *Cell* 1985 they posed the following questions: “(1) *Are there other human or animal diseases that are associated with similarly organized viruses?*; (2) *Is there a precursor to AIDS-associated virus(es) normally present, in latent form in human populations?*; (3) *What triggered in this case the recent spreading of pathogenic derivatives?*” (Wain-Hobson *et al.* 1985).

Phylogenetic methods have been able to demonstrate that HIV originated from multiple interspecies transmissions among simian species which are thought to have originated

in sub-Saharan Africa, where two genetically distinct types of HIV viruses were introduced into the human population (Gao *et al.* 1999). In their review Hahn *et al.* (2000) describe the principal implications of diverse zoonotic infections² in terms of a set of simian immunodeficiency viruses (SIV) found in 26 different species of African nonhuman primates. HIV virus type 1 (HIV-1), which is more virulent and the main cause of HIV infection in pregnant women, as studied in the research reported in this thesis, and type 2 (HIV-2) are the leading causes of AIDS. They are now generally considered to have entered the human population from the cross-species transmission of the simian immunodeficiency virus through zoonosis from primates to humans (Hahn *et al.* 2000). Chimpanzees, *Pan troglodytes*, are found in west equatorial Africa and are considered to be the reservoirs from which HIV-1 was derived (Sharp *et al.* 2011; Hahn *et al.* 2000).

HIV-1 is made up of three groups – M, N and O – which, although responsible for different prevalences, are all capable of causing CD4 T-cell depletion in infected humans and therefore of leading to AIDS (Figure 2-1). The global HIV-1 pandemic is largely due to viral isolates of SIVcpz, which are found in the chimpanzee, belonging to the HIV-1 M group, which was the first to be identified (Sharp *et al.* 2011). The HIV-1 M “primary group” can be classified into nine other subtypes, A–K (Figure 2-1, within red dashed box). The alarming spread of HIV in South Africa and the greater part of the southern African region is caused by HIV-1 subtype C (Wilkinson *et al.* 2015; Sharp 2011). The origin of HIV-2 was first resolved in 1989 (Hirsch *et al.* 1989) when it was found to be closely related to the SIV found in *Cercocebus atys*, sooty mangabeys (SIVsmm), as indicated in Figure 2-1 (blue dashed box). HIV-2, which comprises of groups A and B, is more prevalent than HIV-1 and most common among the human inhabitants of West Africa (Sharp *et al.* 2011).

² Zoonosis is an infectious disease of animals that can be transmitted naturally to humans. Several contemporary diseases originate via viruses, bacteria or fungi transferred to humans through zoonosis, e.g. AIDS, Ebola virus disease, salmonellosis and various forms of influenza.

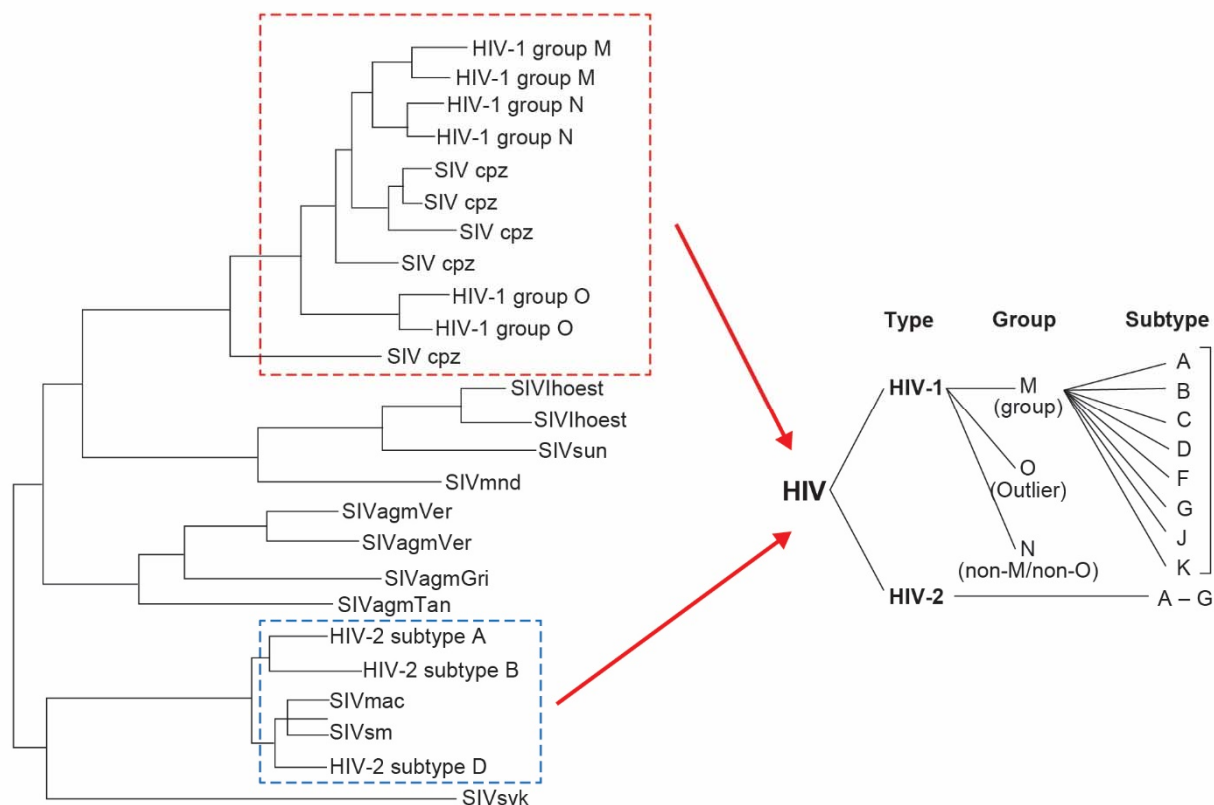


Figure 2–1: Phylogenetic relationships among representative primate lentiviruses. For HIV-1 the group of origin (red dashed box), and for HIV-2 the subtype (blue dashed box), are indicated both on the left and right section of the figure. SIV strains have their host species of origin: cpz, chimpanzee (HIV-1) and mac, macaque; sm, sooty mangabey (HIV-2). (Adapted from Sharp *et al.* 2001; Williams 2012).

2.2 Epidemiology and South African’s brief history

Several research centres, collaborating groups and individual researchers are involved in HIV/AIDS research in South Africa. Some reference will be made in this overview to contributions from these centres and individuals, but it will not be comprehensively covered, as this would require a review in its own right.

In a widely quoted review, Cohen *et al.* (2011) covered critical advances on HIV-1 transmission and acute HIV-1 infection. HIV-1 is primarily a sexually transmitted disease that causes infection through the exposure of mucosal surfaces to the virus (Cohen *et al.* 2011). It is traditionally transmitted through the contact of certain body fluids of HIV-infected individuals during sometimes referred to as ‘unprotected sex’ or

‘unprotected sexual intercourse’. Heterosexual transmission accounts for the majority (85%) of all HIV-1 infections. Since its identification in the 1980s, more than three decades ago, HIV has infected a total of over 70 million people and has claimed more than 35 million lives. Approximately 37 million people (Figure 2-2) are currently living with the disease, 1.8 million of whom are children (less than 15 years of age) (WHO 2015). By the end of 2015, it was reported that 2.1 million people had been infected in that year with HIV globally. Sub-Saharan Africa, a region that bears the brunt of this infection, hosts two-thirds of the global total HIV-infected population; whom 61% of adults living with HIV/AIDS are woman of childbearing age (15–35 years). Consequently, of the reported 25.6 million sub-Saharan people living with HIV, almost 90% of all HIV-positive children live in this region, mainly infected through MTCT (UNAIDS 2015). Within the global context, South Africa, a country heavily affected with illnesses due to HIV/AIDS, has witnessed exceptionally high prevalence rates.

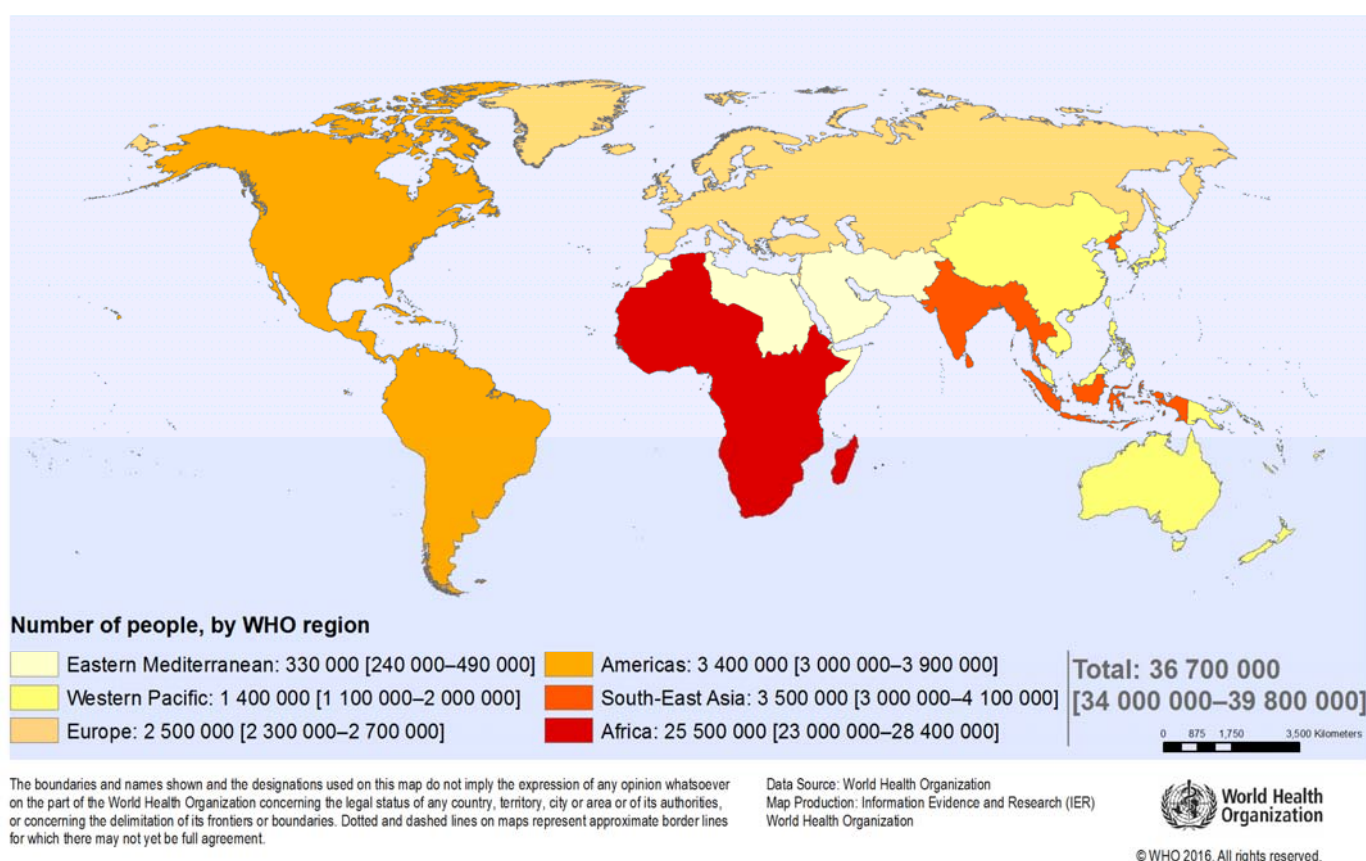


Figure 2–2: Adults and children estimated to be living with HIV in 2015 by WHO region (WHO 2016).

2.2.1 HIV in the South African socio-political context

In 1982, when the first AIDS-related death occurred, the South African socio-political dispensation of *apartheid* (1948–1994) served tragically as the perfect environment where HIV thrived and was ignored over the next decade (Simelela & Venter 2014; Karim & Karim 2002). The exponential growth of the mining sector in the 19th century in South Africa resulted in profound socio-economic changes. The rapid surge of international investments flowing into the country led to an increase in the procurement of cheap black male labour followed by a massive influx of migrants into the country. The implementation of strict legislation restricting native Africans to access to land or to the means of economic production increased the movement of black males into urban areas (Coovadia *et al.* 2009). The poorly paid migrant labour system forced the black male population to work for less than living wages and began the displacement of black men from their families. This extreme segregation made them more susceptible to multiple sexual partners and to establish second families in urban areas. Women who were left in the rural areas also often took other sexual partners, and those in the towns and cities subjected themselves to sex-work in and around the mining compounds (Lurie 2003). These practices increased the risk of sexually transmitted diseases to both migrant workers and to the native black male population.

By 1980 the mining sector had recruited approximately 1.5 million migrant workers (Chirwa 1998). This forced the South African apartheid government to screen Malawian migrants following the increased prevalence of HIV in Malawi (Crush *et al.* 2005; Clark & Worger. 2013). The review by Wilkinson *et al.* (2015) on the history and origin of HIV-1 in South Africa and the greater southern African region describes how, by 1992, the antenatal prevalence in Zimbabwe was 12.5%, in Zambia 9.4% and in South Africa 2%. These high and growing rates indicate the increase in HIV prevalence due to the social, political and economic effects on the disease pattern both in South Africa and in neighbouring countries (Wilkinson *et al.* 2015).

HIV thrived in South Africa under a “*perfect storm*”. The apartheid regime gave little if any attention to the pandemic because it was regarded as a black and gay problem (Karim & Karim 2002). This unattended, uncontrolled increase in the spread of HIV resulted in a subsequent rise in prevalence. In 1990 antenatal HIV prevalence nationally was 0.76%; by 2001 it had reached 24.8% (Department of Health RSA 2001).

2.2.2 Denialism in a post-apartheid health system

Following a period of transition (1990–1994), South Africa entered a new socio-political dispensation as a fully inclusive democracy. Under the leadership of the African National Congress (ANC), the country held its first democratic elections in 1994. With the ANC's focus merely on the removal and amendment of unjust apartheid policies that emasculated the country for many years, AIDS took a backseat in the policies of the new democratic government (Coovadia *et al.* 2009; Karim & Karim 2002). As mentioned above, over the past decade South Africa has experienced a dramatic and rapid growth in the prevalence of HIV. Prevalence has consistently been highest in KwaZulu-Natal where 33.5% of pregnant women are HIV-infected. Five million South Africans, which is more than 5% of the country's population, are now infected with HIV (Lurie *et al.* 2003). Post-apartheid South Africa saw an era of a devastating spread of infection that showed no signs of a plateauing out and which affected mainly women of a child-bearing age (Karim & Karim 2002).

Under the stewardship of President Thabo Mbeki, South Africa was placed under heavy scrutiny for his catastrophic denialism and non-scientific belief that "*HIV did not cause AIDS*" (Schneider *et al.* 2002; Simelela & Venter 2014). This amounted to hundreds of thousands of lost lives as a result of the restriction of a roll-out of PMTCT programmes and antiretroviral treatments (Nattrass 2008). These disastrous effects were reported by Chigwedere *et al.* (2008), who estimated that "*more than 330 000 lives or approximately 2.2 million person-years were lost because a feasible and timely ARV treatment program was not implemented in South Africa*" (Chigwedere *et al.* 2008). Thabo Mbeki and his Minister of Health, Manto Tshabalala-Msimang were gravely misled and held on to unsubstantiated conspiracy theories and strongly believed that antiretrovirals (ARV) were highly toxic (Nattrass 2008, 2013).

2.2.3 The mother-to-child transmission of HIV

Prenatal infection with HIV forms a class of its own in the HIV/AIDS scenario, commonly known as MTCT or vertical transmission (Read *et al.* 1999). In children, MTCT is the major cause of HIV infections, which can occur via various routes; *in utero* (antepartum) during pregnancy, at birth during delivery (intrapartum), and after delivery (postpartum) through breastfeeding (Sripan *et al.* 2015). Prior to 1995, perinatally infected infants had a significantly equal chance of developing AIDS before their third

birthday and an even higher percentage of dying by 10 years of age (Pliner *et al.* 1998). By the end of 2000, substantial progress had been made in the battle against the transmission of HIV-1 from mother to newborns in industrialized countries. The use of antiretroviral therapy (ART) during pregnancy and the neonatal period indicated a remarkable reduction in the rate of MTCT of HIV from 25% to less than 2% (Cooper *et al.* 2002). In 2000 civil society movements by researchers and activists in South Africa took a stand and appealed to the Constitutional Court to contest the government's inaction in working with the Global Fund and PEPFAR in providing life-saving treatment and public ART programmes (Katz *et al.* 2013).

As mentioned earlier (Chapter 1), in 2011 PMTCT interventions were implemented by WHO, along with several initiatives taken by UNAIDS to set a new goal to virtually eliminate new paediatric HIV infections by 2015 "... *elimination of new infections by 2015....*". In a recent follow-up interview titled "15 million people in conversation with Michel Sidibé, regarding the progress of this goal. The Executive Director of UNAIDS Michel Sidibé's response was as follows: "*We launched the Global Plan with PEPFAR in 2011.... The results—73% of all pregnant women living with HIV have access to antiretroviral medicines and we reduced new HIV children infections by 58%. We also improved the quality of medicines for women and children*" (UNAIDS 2015).

While studies have identified that most transmission is believed to occur before and during delivery, MTCT risks vary with whether the affected population is breastfeeding or non-breastfeeding (Newell 2001). The geographical settings of whether the population is in a rich or poor country also plays a role and largely determines whether access to a full range of ARV prophylaxis is possible as part of antenatal care. Many women in sub-Saharan Africa never benefit from such access and therefore the relevant programmes are rarely initiated (De Cock *et al.* 2000). Despite the enormous efforts of the global community and public health sector towards the elimination of new HIV infections among children by 2015, HIV infection still largely affects infants and remains a paediatric pandemic even with the substantial reductions in MTCT. In the 2015 WHO report, "*When to start antiretroviral therapy guideline*", it is stated that: "*Adolescents and young women living with HIV-1 face unique challenges in preventing the transmission of HIV-1 to their children and attending to their own health needs, including poor access to reproductive health services, poor uptake of testing and poor retention in care*" (WHO 2015). South Africa currently runs one of the most successful PMTCT programmes, which is funded by the government with more than 3 million people on ARV (UNAIDS 2015). Risks associated with PMTCT include maternal,

obstetrical and neonatal factors; although their timing is still speculative, the risks are linked to maternal HIV-health status, neonatal feeding and fetal exposure to infected maternal body fluids.

2.3 Mother-to-child transmission (MTCT) and risk factors

2.3.1 Timing of MTCT HIV transmission

The exact timing and mechanisms of the viral transmission from mother to her child are still largely unknown and remain under investigation (Petropoulou *et al.* 2006). Knowledge about the timing of transmission is crucial. Its understanding would provide more refined treatment strategies to help prevent MTCT viral infection and recommendations of when to start them (De Cock *et al.* 2000; Bryson *et al.* 1992; Simonon *et al.* 1994). It appears to be hard to provide evidence of the exact transmission timing as pointed out by many researchers. However, a review by Kourtis *et al.* (2006) discusses the implications of a hypothesis based on recent research findings regarding viral latency, that the initial time of fetal viral exposure might not be when infection is established and the subsequent implications for its prevention. To deduce the timing of MTCT of HIV Kourtis *et al.* (2006) analysed the data from 18 major clinical trials that investigated PMTCT with antiretroviral regimes. They provide a range of estimates that are of practical use in the planning and evaluation of intervention trials designed to interrupt transmission during specific risk periods (Table 2-1). Kourtis and co-workers assert that, for non-breastfeeding populations, half of the viral infections takes place at the end of pregnancy and towards the time of labour. This conclusion is similar to results reported in a much older study by Bertolli *et al.* in 1996, in a population studied in Kinshasa, Zaire.

Table 2–1: Estimation of timing of MTCT in non-breastfeeding and breastfeeding populations (adapted from Kourtis *et al.* 2006).

Population	Non-breastfeeding	Breastfeeding 18–24 months	Breastfeeding 6 months
Before 14 weeks (%)	3	2	4
14–36 weeks (%)	17	10	13
36 weeks through labour	50	29	39
Delivery	30	20	26
Postpartum	n/a	39	18
Total MTCT (%)	25	41	31

n/a, not applicable.

It is estimated that transmission risks during breastfeeding are higher during early lactation periods (< 6 months), due to the increased viral loads in colostrum and just as high in women who breastfeed for prolonged periods (De Cock *et al.* 2000). For the breastfeeding population, the greatest risk of viral transmission takes place in the postnatal period (Kourtis *et al.* 2006). Breastfeeding risks, however, are dependent on the maternal CD4 count, the duration of feeding, maternal viral loads and mixed feedings (both breast milk and breast milk substitute such as commercial infant formula). The model presented in Table 2-1 suggests that if the time of breastfeeding is shortened from 18–24 months to 6 months, transmission risks during this period could drop significantly, further reducing the total rate of MTCT. Improved knowledge of timing regarding transmission through breastfeeding is crucial and would further improve on the guidelines for optimal feeding practices (Kourtis *et al.* 2006).

2.3.2 Maternal factors

Risk factors for vertical MTCT of HIV, which increase the likelihood of infection, are well defined. They largely relate to the mother's immune status, measured as the peripheral blood's viral load or as clinical and immunological markers (Tolle & Dewey 2010). Maternal HIV RNA levels at delivery have been described in various studies, and are reported as a strong and consistent predictor of the risk of transmission (Mofenson *et al.* 1999; Burns *et al.* 1997). Increased maternal viral loads of plasma RNA levels of more than 100 000 copies/mL have been linked to higher transmission rates (Petropoulou *et al.* 2006). Effective antiretroviral treatments, such as combination antiretrovirals (cART), have been known to reduce the viral load and therefore to

increase its efficacy in preventing MTCT. However, there have been occasional reports, such as that of the AIDS Clinical Trial Group protocol 076, which found that transmission occurred over a range of maternal RNA levels, even in mothers whose viral loads were too low to be counted (Sperling *et al.* 1996). Therefore, it seems that when it comes to maternal viral factors there is no viral load threshold below which there is no risk of transmission of perinatal infection. Although the maternal viral load is an important immunological parameter, significant in predicting perinatal transmission, it is not the only one (Newell 2001). Lower maternal CD4 cell counts have been reported to be associated with higher risks for MTCT. Antiretroviral regimes that reduce HIV RNA to lower levels than 1000 copies/mL increase maternal CD4 counts and are associated with reduced risks of perinatal transmissions as well as improved health of the mother (Burns *et al.* 1997).

2.3.3 Obstetrical factors

Chorioamnionitis, an inflammation of the lining of the chorion-amnion space, is reported to slightly increase the risks for MTCT (Tolle & Dewey 2010). Placental inflammation in women who present with low levels of HIV in their blood could increase the risk of transmission, owing to the compromised barrier that separates the mother's blood from the baby along with other secretions (Mofenson *et al.* 1999). Women who experienced premature rupture of membranes, such as those reported in the randomized Perinatal HIV Prevention Trial-1, which enrolled 1437 women, had an estimated risk of intrapartum transmission of the virus of 7.7% if labour was induced and 6.9% if it was not (Jourdain *et al.* 2007). The duration of the ruptured membranes in women who experienced prolonged labour of more than 4 hours was reported to be related to an increased transmission risk factor of about 2% for every hour. Mothers with relatively low CD4 cell counts who encounter prolonged rupture of their membranes are at an even greater risk.

2.4 Prevention of mother-to-child transmission (PMTCT)

2.4.1 Setting the stage

“In pregnant women with mildly symptomatic HIV disease and no prior treatment with antiretroviral drugs during pregnancy, a regimen consisting of zidovudine given antepartum and intrapartum to the mother and to the newborn for six weeks reduced the risk of maternal-infant transmission by approximately two thirds.” (Connor et al. 1994)

This conclusion from a 1994 American/French cohort study, designated as PACTG-076, set forth a paradigm shift and paved the way forward on the prevention of vertical HIV transmission from mother to child.

A cohort of 477 HIV-positive pregnant women were administered zidovudine intravenously between 14–34 gestational weeks until delivery, with their newborns placed on the treatment for 6 weeks. The success in their findings yielded a 67% reduction in the rate of maternal transmissions. Although the WHO’s ART guidelines have advanced since they were first issued in 2002, the PACTG-076 trial remains the backbone of ARV regimens provided as treatment. This treatment has averted over 1.2 million new HIV infections among children and prevented an estimated 4.2 million deaths in resource-poor settings since 2009 (UNAIDS 2015). PMTCT of HIV is generally regarded as one of public health’s greatest successes.

Many studies, including this PhD, have reported on the safety and consequences of using ARV agents during pregnancy, especially since there is an increasing cohort of children born HIV-exposed but uninfected. Women who are successfully initiated into an ART programme are often prescribed highly active antiretroviral therapy (HAART). To date, there is no cure for HIV so that lifelong treatment is necessary and the best option for long-lasting suppression of viral replication and HIV-associated inflammation. Table 2-2 presents the classes of the most extensively used ARVs globally over the years. These include: nucleoside/nucleotide analogue reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs).

**Table 2–2: Antiretroviral drugs used on pregnant HIV-infected women
(Adapted from Simon *et al.* 2006).**

FDA-approved antiretroviral drugs: Generic name (abbreviation) Trade name					
Entry	Reverse transcriptase inhibitors [NRTIs ^{a1} & NNRTIs ^{a2}]			Protease inhibitors [PI ^b]	
		Nucleoside	Nucleotide	Non-nucleoside	
Single compound drugs	Enfuvirtide (T20) <i>Fuzeon</i>	Abacavir (ABC) <i>Ziagen</i>	Tenofovir (TDF) <i>Vired</i>	Efavirenz (EFV) <i>Sustiva</i>	Fosamprenavir (FPV) <i>Lexiva</i>
		Didanosine (ddl) <i>Videx/Videx EC</i>		Nevirapine (NVP) <i>Viramune</i>	Atazanavir (ATV) <i>Reyataz</i>
		Emtricitabine (FTC) <i>Emtriva</i>			Darunavir (DRV) <i>Preszista</i>
		Lamivudine (3TC) <i>Epivir</i>			Indinavir (IDV) <i>Crixivan</i>
		Stravudine (d4T) <i>Zerit</i>			Nelfinavir (NFV) <i>Viracept</i>
		Zidovudine (AZT/ZDV) <i>Retrovir</i>			Ritonavir (RTV) <i>Norvir</i>
					Saquinavir (SQV) <i>Invirase</i>
					Tripanavir (TPV) <i>Aptivus</i>
					Lopinavir/Ritonavir (LPV/r) <i>Kaletra</i>
Fixed-dose combination drugs (cART)		Abacavir/Lamivudine (ABC/3TC) <i>Epzicom</i>			
		Zidovudine/Lamivudine (ZDV/3TC) <i>Combivir</i>			
		Tenofovir/Emtricitabine (TDF/FTC) <i>Truvada</i>			
		Abacavir/Lamivudine/Zidovudine (ABC /3TC/ZDV) <i>Trizavir</i>			
		Tenofovir/Emtricitabine/Efavirenz (TDF/FTC/EFV) <i>Atripla</i>			

Note: Drugs belong to five different classes that target three different viral steps (entry, reverse transcription or protease).

^{a1} NRTIs are recommended for use as part of combination regimens, usually including two NRTIs with either an NNRTI or one or more PIs. Use of single or dual NRTIs alone is not recommended for treatment of HIV infection.

^{a2} NNRTIs are recommended for use in combination regimens with 2 NRTI drugs. Hypersensitivity reactions, including hepatic toxicity and rash, are more common in women; it is unclear if these reactions increase in pregnancy.

^b PIs are recommended for use in combination regimens with 2 NRTI drugs. Hyperglycaemia, new onset or exacerbation of diabetes mellitus, and diabetic ketoacidosis reported with PI use; it is unclear if pregnancy increases risk.

2.4.2 WHO PMTCT recommendations and strategy for pregnant women

The most critical step prior to administration of appropriate antiretroviral drugs to HIV-infected individuals is to obtain relevant clinical information on them. This was particularly emphasized in the WHO guidelines regarding the importance of antenatal HIV-testing, offered as a standard to every pregnant woman, to identify her status. But fundamentally, to start her on antiretrovirals which will improve her own health therefore providing appropriate care (Doherty *et al.* 2013, WHO 2015). It is essential that pregnant women undergo repeat testing throughout their pregnancy. Women who test negative at their first antenatal test have been found to test positive at a later stage; the acquisition of HIV in these women can occur during pregnancy or postpartum as indicated in various studies (Drake *et al.* 2014; Moodley *et al.* 2009, 2011). Failure to retest could lead to missing women with late incidence of infection, increasing their risk of MTCT and resulting in fewer chances of them ever being initiated into an ART programme.

The 2011 launched call for 15 million people to be on treatment by 2015, saw an increase in the focal shift on prevention of perinatal infection rather than ignoring the concerns for the health and lives of the mothers living with HIV (UNAIDS 2016). This was owing to the crucial strides made in 2010 by the WHO, which announced PMTCT guidelines emphasizing the importance of providing more efficacious regimens, coherent in preventing MTCT during pregnancy, labour and delivery, postpartum and throughout breastfeeding (WHO 2010). These guidelines were based on two approaches: (1) **Lifelong antiretroviral therapy (ART)** — *“HIV-infected woman in need of treatment for their own health, which is also safe and effective in reducing MTCT”* [CD4 count $\leq$ 350 cells/mm³ such as WHO stage 3 or 4]. (2) **Antiretroviral (ARV) prophylaxis** — *“to prevent MTCT during pregnancy, delivery and breastfeeding for HIV-infected women not in need of treatment”* [CD4 count $>$ 350 cells/mm³] (WHO 2010).

Prior to the WHO's 2012 consolidated guidelines, mothers in resource-poor settings were assigned Option A (CD4 counts $\leq$ 350 cells/mm³, ART for life; CD4 counts $>$ 350 cells/mm³, ARV prophylaxis as early as 14 weeks gestation) largely owing to its affordability. Now updated, PMTCT Option B (CD4 counts $\leq$ 350 cells/mm³, ART for life; CD4 counts $>$ 350 cells/mm³, Triple ARV starting as early as 14 weeks) and B+ (long-life triple ARV regardless of CD4 count) are the preferable options because of their greater efficiency. This is because these two options offer benefits beyond PMTCT and address the woman's health by offering her treatment earlier. The long-

term benefits and potential risks of the B+ approach for mother and child are still debated issues, as discussed below.

2.5 Long-term consequences of HIV-ART exposure

A review by Ahmed, Kim and Abrams (2013) gives a detailed assessment of the risks and benefits of Option B+, following the WHO update on lifelong Option B+ ARV for all pregnant and breastfeeding women. The concerns conveyed by Ahmed and his colleagues (Ahmed *et al.* 2013) regarding increased exposure to ART for a longer duration, which in some women can be as early as conception, is that it could lead to risks for the fetus such as birth defects and toxic consequences. They finally recommend that, as national programmes increasingly adopt this approach, “it will be critical to carefully assess both short-term and long-term infant outcomes” – a key aspect that motivated the present study. The WHO followed suit. In their 2015 guidelines: “When to start antiretroviral therapy and on pre-exposure prophylaxis for HIV”, the WHO states that “*the risks for pregnant and breastfeeding woman are similar to those for non-pregnant adults. However, a further potential risk to the fetus may arise from exposure of ARV drugs given as early as gestational period*” [underlined: my emphasis]. It is commonly stated in most of the WHO guidelines that no ARV drug is without risk to the fetus. Moreover, although ART in PMTCT during pregnancy and breastfeeding has indicated clear public benefits for both maternal health and reduced MTCT, it should be noted that the potential long-term effects of *in utero* and fetal exposure to these maternal drugs are still poorly understood (WHO 2015).

Another group of investigators who assessed the safety and effects of the widely recommended B+ PMTCT regimens were Tsague & Abrams (2014). They comment on the first-line drugs tenofovir (TDF) + lamivudine (3TC)/efavirenz (FTC) + emtricitabine (EFV) (see Table 2-2), recommended as part of the lifelong treatment for women. Although data on both emtricitabine and tenofovir suggest no increase in poor pregnancy outcomes, research on their use in pregnant women is limited. However, there have been studies that warn of the controversial use of efavirenz during pregnancy, given its teratogenic effects observed in animal studies on *in utero* exposure to the regimen along with the neural tube defects presenting in the first trimester (De Santis *et al.* 2002; Saitoh *et al.* 2005). Tenofovir has been linked to renal damage and bone density loss in adults although prenatal studies exploring *in utero* infant TDF (Table 2-2) exposure have advocated their use with no links to poor pregnancy outcomes (Colbers *et al.* 2013). However, since the long-term effects of tenofovir in infants is unknown, studies are warranted for further longitudinal

investigation to evaluate the outcome to its exposure (Tsague & Abrams 2014; Ahmed *et al.* 2013).

With the increase in PMTCT ARV accessibility through WHO initiatives, more HIV-infected mothers will give birth to HIV-negative children. This, however, gives rise to a cohort of infants who are born exposed to the virus (Kellerman *et al.* 2013). In South Africa, there is an increase in a cohort of HIV uninfected children born to HIV-positive mothers; this region with an additional 1.2 million infection of childbearing women between 2010–2015, therefore has an increasing prevalence of uninfected yet HIV-exposed children (UNAIDS 2015). Slogrove *et al.* (2016) reviewed the clinical consequences of HIV infection for infants and investigated if exposure to the virus in HIV-uninfected children (HEU) plays a direct role in their risks for morbidity and mortality relative to children who are HIV-unexposed (HE). Their review included 22 studies on well-defined HEU children from HE cases, evaluating all-cause mortality, and all-cause hospitalization along with the identified infectious morbidities. From these studies Slogrove and co-workers (2016) comment that more research on interventions devoted to improving health outcomes in HEU is needed. They conclude that *“HEU infant health and well-being, beyond avoiding HIV-infection, deserves a more prominent position in the local and international HIV research agendas.”* (Slogrove *et al.* 2016).

2.6 Metabolic studies of HIV and ART exposure

2.6.1 Metabolic and mitochondrial indicators of HIV and ART exposure

Regarding metabolic and mitochondrial indicators of HIV and ART exposure, my focus is on potential consequences of ART for PMTCT, as the effects were already noted in a pioneering study on a cohort of HIV-ART-exposed neonates (Connor *et al.* 1994). These authors reported that minimal short-term toxic effects were observed in the cases studied, but did not fail to note that the level of haemoglobin at birth was significantly lower in the ART-exposed infants than in the control group, although haemoglobin concentrations returned to normal 12 weeks after birth. Subsequent studies also indicated that exposure to prophylactic zidovudine for PMTCT is associated with a usually mild and reversible anaemia in infants, but rarely occurs as severe (Sperling *et al.* 1998). These very early observations clearly alert us to the possible toxic effects of ART.

Nearly a decade later, a review by Thorne and Newell addresses the adverse effects of antiretroviral prophylaxis in PMTCT, highlighting several examples. They discuss pregnancy-related effects such as preeclampsia, hyperglycaemia, and mitochondrial toxicity as the main consequences of ART and refer to other, less severe consequences of *in utero* neonatal exposure to ARV (Thorne & Newell 2007). HAART regimens have displayed several class-specific side effects and complications. Dyslipidemia has been observed in HIV-infected patients taking ritonavir, in a PI-based regimen (Mulligan *et al.* 2000). Metabolic lipid disorders in HIV-infected children are, however, relatively scarce as compared to the prevalence seen in HIV-infected adults (McComsey *et al.* 2004).

Recent studies have reported that ARV could have adverse effects regardless of the health status of HIV-infected individuals. Nucleoside analogue and PI-causing mitochondrial toxicity and various linked metabolic complications are well recognized in studies in non-pregnant women, as reviewed by Kirmse, Baumgart and Rakhmanina (2013b). HAART has significantly improved the immunological status of infected mothers and is more effective than mono- or dual-therapy. Although the HAART benefits in reducing MTCT are great, they warn that the treatment exposes the fetus to potential chemotoxins (Mas *et al.* 2012).

HIV and all three major classes of antiretroviral regimes, NRTI, NNRTI and PI (see Table 2-2), have been known to induce mitochondrial depletion and metabolic dysfunction (Kirmse *et al.* 2013b). This is represented schematically in Figure 2-3.

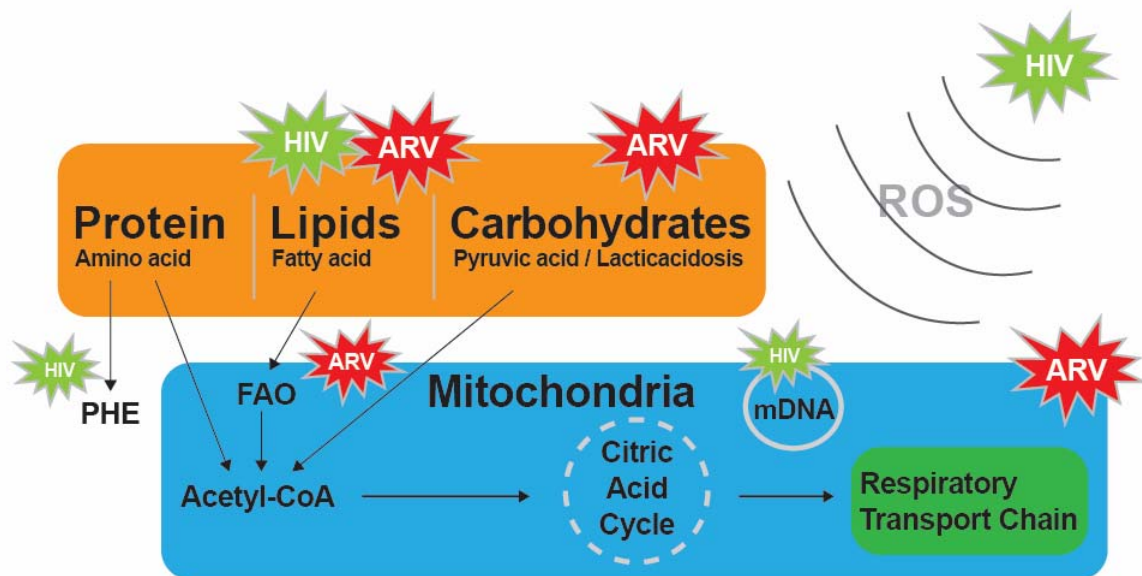


Figure 2–3 A conceptual representation of the effect of HIV and ART on mitochondrial metabolism. Antiretroviral (ARV) exposure is shown to affect lipid (dyslipidemia) and carbohydrate (lactic acidosis) metabolism, fatty acid oxidation (FAO) that lead to the formation of acetyl-CoA, and mitochondrial (oxidative phosphorylation) metabolism. HIV infection itself is primarily associated with a variety of effects on mitochondrial DNA (mtDNA), with increased generation of reactive oxygen species (ROS), dysfunctional lipid metabolism and some minor effects on phenylalanine (PHE) as well as metabolism of some other amino acids. (This diagram is an elaboration of Figure 1 in Kirmse *et al.* 2013b with permission from Kirmse *et al.* 2013b).

- Mitochondrial DNA

Following the first decade of NRTI therapy against HIV infection, several adverse effects were reported and seemed to be the consequence of one common mechanism: “a decreased mitochondrial energy-generating capacity.” (Brinkman *et al.* 1998). Acquired mitochondrial toxicity in patients taking ART occurs due to two key effects: incorporation of the ART nucleoside analogues into mitochondrial DNA (mtDNA) and/or inhibition of mtDNA replication. However, the effect of ART on mtDNA is a complex phenomenon. The observed AZT (see Table 2-2) toxicity in vivo seemed to be through the combined effect of moderately efficient incorporation, inefficient ART removal resulting in persistence in mtDNA, and inhibition of polymerase γ (pol γ) proof

reading causing an increase in mtDNA mutations (Lim & Copeland. 2001). Early in the history of HIV-ART it was cautioned that viral DNA can cross the placenta, can interact with human DNA and that fetal tolerance of these drugs used during pregnancy should be carefully assessed, especially because there can be mitochondrial dysfunction in HIV-exposed uninfected children treated *in utero* and postpartum with zidovudine (Blanche *et al.* 1999). “*There is no obvious reason why the fetus and newborn, often exposed for many months, should escape any such effects.*” (Blanche *et al.* 2006). In their early research, Blanche *et al.* also reported mitochondrial dysfunction in HIV-exposed uninfected children exposed *in utero* and postpartum to NRTI zidovudine. Similarly, in a study by Poirier *et al.* (2003), who compared kids born to HIV-positive mothers who were exposed *in utero* to zidovudine along with newborns who were born to HIV-negative mothers, found that mitochondrial dysfunction had a higher incidence in HIV-exposed uninfected infants than the general population. It was further noted that fetal exposure to multiple NRTIs resulted in higher dysfunctional risks than exposure to monotherapy (Poirier *et al.* 2003). In a mice model study, Chan *et al.* (2006) discovered that their exposure to a combination of zidovudine and lamivudine, a two-drug NRTI, resulted in even greater mitochondrial DNA damage. Depletion of mtDNA following NRTI exposure is now generally accepted, but the implications of these findings are not fully understood. Subclinical mitochondrial toxicity is likely to contribute to some of the longer-lasting biochemical alterations reported in ART-exposed infants (Afran *et al.* 2014). Lactic acidosis and lactic acidemia, my next focus, are important biomarkers of metabolic dysregulation due to congenital abnormalities in carbohydrate metabolism (see Figure 2-3), but are also known to be caused by certain drugs or toxins, and were among the first indicators of ART toxicity in infants.

- Carbohydrate metabolism and lactic acidosis

Lactic acidemia and lactic acidosis were noticed in some ART-exposed infants, as reported by Giaquinto *et al.* (2001). Their study included 20 infants who had experienced varying prenatal or perinatal exposure to ARV compared with a control group of 36 HIV-infected children who had either received ARV therapy or were not so treated. Seventeen out of these 20 children (85%) at least once showed a lactic acid level exceeding 2.5 mmol/L, but elevated lactic acid levels returned to normal during follow-up treatment and none of the children developed any clinical signs or symptoms consistent with mitochondrial damage. However, individual cases of severe lactic acidosis occurred: Scalfaro *et al.* (1998) reported on a case of severe lactic acidosis in a neonate from a zidovudine-treated mother during pregnancy and receiving postnatal zidovudine, but which resolved rapidly following discontinuation of the

zidovudine treatment. The issue of risk for neonates following *in utero* exposure was specifically addressed by Noguera and co-workers (2004). The aim of their study was to determine the prevalence, clinical consequences, evolution, and risk factors for hyperlactaemia (HLA) in their cohort of HIV-uninfected infants who were exposed to ARV during gestation, labour, and/or the neonatal period. They reported the remarkable observation that HLA is higher in HIV-uninfected infants who are exposed perinatally to ARV than in HIV-infected paediatric patients on chronic HAART, and hypothesized “*that the higher energy requirements of a developing organism in the first months of life may make these patients more vulnerable to NRTI toxicity*” (Noguera *et al.* 2004). From their experience on HIV-infected paediatric patients, they conclude that only a younger age at the beginning of ART was found to be a statistically significant risk factor to develop HLA.

- Dyslipidemia

Associated with increased lactic acid, keto acids (for example, 3-hydroxybutyric acid and acetoacetic acid) are produced when carbohydrate or fat is incompletely oxidized. A link has been proposed between maternal lipid and cholesterol abnormalities during pregnancy with a higher risk of preterm delivery (Mudd *et al.* 2012). Moreover, perturbed lipid profiles were already noticed in early studies of HIV-infection and in AIDS patients: plasma triglycerides and free fatty acids were increased in AIDS, whereas plasma cholesterol, high-density lipoprotein (HDL) cholesterol, apolipoprotein-A-1 (Apo-A-1), low-density lipoprotein (LDL) cholesterol, and Apo-B-100 levels were decreased (Grunfeld *et al.* 1992). The underlying mechanisms responsible with regard to ARV-associated dyslipidemia still remain obscure, although it is speculated that they are multifactorial (Kirmse *et al.* 2013b). A recent review on dyslipidemia related to ART (Estrada & Portilla 2010) indicates that many steps of lipid and lipoprotein metabolism are involved, including increased lipolysis, altered receptor expression, regulatory element degradation and very-low-density lipoprotein catabolism (Estrada & Portilla 2010). It is also suggested that HIV infection itself and alterations in glucose homeostasis may likewise add biochemical pressure to the systems metabolizing lipids.

- Amino acid metabolism

Amino acids, catabolic products from proteins, the third major group of macromolecules involved in intermediary metabolism, might likewise be affected by the dual effect of HIV and ART. Changes in the metabolism of the amino acids tryptophan and phenylalanine have been associated with HIV disease and were

proposed to be of predictive value following a study of macaques infected with simian immunodeficiency virus (Drewes *et al.* 2015). Both amino acids are precursors for neurotransmitter biosynthesis and may provide a link to the development of neurocognitive impairments seen in HIV-1-infected patients during ART and in AIDS (Gostner *et al.* 2015). Although an increased ratio of phenylalanine to tyrosine has been reported in association with HIV infection, but which could be normalized following ART (Zangerle *et al.* 2010), no convincing indications have yet been found on the disturbed catabolism of phenylalanine, tyrosine, tryptophan and keto acids derived from other amino acids as well as of acetyl-CoA degradation in mitochondria (see Figure 2-3).

- Oxidative phosphorylation

It is known biochemically that genetically inherited defects in mtDNA or in nuclear genes that encode the oxidative phosphorylation system (OXPHOS), can cause an impaired OXPHOS, leading to a variety of clinical disorders due to a dysfunctional ATP system (Greaves *et al.* 2012; Brinkman *et al.* 1999). Skeletal muscle, heart, liver, the nervous system and kidneys, among others, are organs and tissues with higher metabolic energy requirements are at most risk, and present with clinical manifestations when an energy deficit arises (Greaves *et al.* 2012). An interruption in the oxidative phosphorylation system as a result of minimum energy levels could lead to a disturbed redox state, which includes dysfunction of the Krebs cycle. This disturbance in OXPHOS places pressure on the Krebs cycle, shifting pyruvate towards the production of lactate and leading to lactic acidosis, as indicated above. Similarly, a ketogenic occurrence of increased acetyl-CoA is observed (Brinkman *et al.* 1999). Mitochondrial toxicities associated with PMTCT ART exposure have been explored in clinical and experimental studies (Blanche *et al.* 1999; Dagan *et al.* 2002; Lewis *et al.* 2003).

The enzymatic stages of mammalian mitochondrial β -oxidation were elucidated some 50 years ago, which was followed by the discovery of a linked multifunctional membrane-associated acyl-CoA dehydrogenase and a carnitine palmitoyltransferase system. Inherited deficiencies in the enzymes of the β -oxidation system and of the import, processing and assembly of these enzymes within the mitochondria, has led to a greater understanding of the molecular mechanisms of these processes. In a study that involved abnormal newborn screens and acylcarnitine analysis, Kirmse and co-workers (2013a) hypothesized the following: Children born to HIV-positive mothers exposed to HIV-ART would encounter a disruption in their metabolism dependent on a functional OXPHOS, and that this will result in a larger occurrence of positive

metabolic newborn screens for ART-exposed infants against the general population. Their study involved the analysis of data of abnormal newborn metabolic screens in HIV-positive and HIV-negative neonates. Acylcarnitine levels in plasma of 16 ARV-exposed compared to 14 ARV-unexposed infants exposed to HIV were measured by inspecting fatty acid oxidation (Kirmse *et al.* 2013a). Results from the acylcarnitine analysis indicated that ARV-exposed neonates were more likely than the unexposed to have fatty acid oxidation dysfunction, which could lead to disruption in intermediary energy metabolism. An earlier study whose objective was to detect possible carnitine deficiency in HIV-infected children on antiretroviral therapy, reported significantly low levels of serum-free carnitine, acylcarnitine, methionine and lysine compared to reference values of children of similar ages ($P < 0.0001$; Student's t-test) (Vilaseca *et al.* 2003).

- Reactive oxygen species and oxidative stress

A review by Alfadda *et al.* (2012) indicates that work published on reactive oxygen species (ROS), indicators of oxidative stress, dates all the way back to 1945. Between then and the 2012 review article, 117,000 articles in English and 12,000 review articles have been produced on ROS and/or oxidative stress. A wide range of these papers associate ROS with diseases such as cancer, insulin resistance, diabetes mellitus, and cardiovascular diseases. Others link ROS to physiological processes and essential protective mechanisms (Alfadda *et al.* 2012). This large interest in ROS indicates the critical role ROS plays in the regulation of homeostasis in both health and disease. However, in infection the role of ROS in the immune system is critically involved in both the innate and acquired responses. Once exposed to a foreign disease-causing agent, the first line of defence against the foreign pathogen is an explosive release of large amounts of ROS, as part of the oxidative activated phagocytes present at the site of local inflammation (Alfadda *et al.* 2012). ART-induced oxidative stress in HIV exposed and infected patients has been widely researched and considered a cornerstone discovery in redox biology (Ivanov *et al.* 2016). Mandas and associates (2009) analysed serum oxidant and antioxidant levels in patients who were HIV-positive and HAART-treated compared to HIV-positive and HAART-untreated in addition to being HIV-negative. Their investigation revealed higher serum oxidant levels in HIV-infected subjects than controls, as expected. Interestingly, a higher oxidative status was observed in patients who rigorously followed their HAART therapy as compared to those who adhered poorly to their treatment regimen. This suggests that HAART treatment made the oxidative stress even greater (Mandas *et al.* 2009). HIV incites oxidative stress, which leads to an increase in ROS production through the

deregulation of oxidative stress pathways. This causes further mitochondrial dysfunction.

Oxidative stress in HIV has been linked to a plethora of virus-associated pathologies. In the central nervous system, neurotoxicity and dementia have been associated with direct consequences of an HIV infection (Ivanov *et al.* 2016). Proteins Gp120, Tat and Vpr play a substantial role in central nervous system (CNS) pathology, whether it be directly or indirectly. However, if directly involved these proteins are implicated in ROS production and this causes a decrease in cell viability and a progressive exhaustion of the antioxidant defence (Shah *et al.* 2013). Indirectly, they release macrophage, microglia and astrocyte toxins, this action of ROS leads to neuronal injury (Kaul *et al.* 2001).

A stressful intrauterine environment in HIV-ART-exposed neonates, likely due to the *in utero* metabolomic stress can cause an increase in ROS production (Kirmse *et al.* 2013b). Oxidative stress in HIV-associated lung pathology leads to a decrease in expression of the tight junction receptors, leading to a decrease in lung function such as damaged epithelium, thereby increasing the lungs' susceptibility to infections (Beck *et al.* 2001). Ivanov and colleagues (2016) further summarize current data and knowledge on oxidative stress markers and their association with HIV infection in their review. Ivanov *et al.* (2016) analysed the mechanisms through which ROS are triggered by the virus and describe the status of various defence mechanisms of the infected host cell. They analysed various ROS-associated pathologies driven by HIV oxidative stress and point out that "there are still notable gaps in the field that can become targets for future studies.... Continuation of these studies would contribute to the development of efficient antiretroviral treatments and vaccines" (Ivanov *et al.* 2016).

To conclude:

The effect of HIV and ART on metabolic and mitochondrial processes has been and remains a high priority in HIV research. I do not provide an exhaustive coverage of the published material, but what is presented here is sufficient to define the key trends in research on HIV-ART exposure. From this overview, a number of key conclusions can be drawn:

- (1) The metabolic consequences of HIV and ART exposure are clearly multifactorial (Kirmse *et al.* 2013b) – no single, well-defined metabolite could be identified as a biomarker for either HIV or ART, in comparison to metabolites that are evident biomarkers for the monogenic inherited metabolic disorders, e.g. N-isovalerylglycine for isovaleric aciduria.
- (2) Antiretroviral drugs, unlike most other pharmaceutical agents, have become widely prescribed in pregnancy without due consideration of potential risk and safety issues. The assumed potential risks associated with ART exposure for pregnant woman and their infants depends on the drug, its dosage and duration of the exposure have not yet been optimally determined (Ahmed *et al.* 2013).
- (3) Although antiretroviral drugs are generally considered to be safe and the global response to care for infected and exposed children has matured, individual infants highly affected by ART exposure were and are periodically reported. Surveillance on the long-term effect of ART on the growing cohort of children born exposed to ART but are HIV uninfected should continue (Kellerman *et al.* 2013).

2.6.2 Metabolomic studies of HIV

Details on metabolomics technology are presented in Chapter 3. This chapter will introduce metabolomics terminology, followed by a review of metabolomics studies on HIV, ART and PMTCT. This account leads to and concludes with the aims of and biological questions addressed in the present study.

The term “*metabolome*” was first used by Olivier *et al.* (1998) to describe the total set of metabolites synthesized by organisms. The term was fashioned analogous to the “genome” (total DNA content of the genome of an organism) and “*proteome*” (total protein content of an organism). In biochemical terminology, the metabolome refers to the quantitative complement of all of the low molecular weight molecules [i.e. metabolites with molecular weight less than 1 000 daltons] present in cells in a particular physiological or developmental state (Olivier *et al.* 1998). Analysing biological systems by monitoring the metabolome has been shown to be an effective tool for disease diagnosis and characterization of biological pathways, thereby capturing the status of the diverse biochemical pathways at a particular moment in time (Wikoff *et al.* 2007).

Initially, two terms were defined to describe the scientific approach for investigations of the metabolome. Metabonomics was introduced by Jeremy Nicholson of Imperial College London (Nicholson *et al.* 1999). He proposed “a new NMR-based ‘metabonomics’ approach... that is aimed at the augmentation and complementation of the information provided by means of the genetic and proteomic responses to xenobiotic exposure”. Nicholson specifically defined metabonomics as “*the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification*” (Nicholson *et al.* 1999). Metabolomics was introduced somewhat later by Oliver Fiehn of the Max Planck Institute of Molecular Plant Physiology, Potsdam (Fiehn 2002). Fiehn highlighted the need for “*a comprehensive analysis in which all the metabolites of a biological system are identified and quantified.... Since such an approach reveals the metabolome of the biological system under study, this approach should be called metabolomics*” (Fiehn 2002). Initially, there was some confusion over the double terminology, but at present the two terms are used interchangeably, with the tendency in the literature towards the term metabolomics, which will be used in this thesis. The implementation of metabolomic techniques includes two complementary approaches: ‘metabolite profiling’ and ‘metabolomic fingerprinting’. Metabolomic profiling typically seeks “*to identify and quantify a particular set of metabolites belonging to a specific pathway or class of compounds*” (Carraro *et al.* 2009). This approach seeks to answer investigative questions with no intention of finding unknown metabolites (Dettmer *et al.* 2007). However, metabolic fingerprinting is a more comprehensive approach and is more likely to be considered as a genuine “*omic*” approach and a true representative of metabolomics. In metabolomic fingerprinting no single metabolite or group is sought out, instead it regards the entire profile as a ‘fingerprint’ and tries to distinguish metabolic characteristics that discriminate between groups of subjects. This includes biochemical patterns related to certain pathological conditions, exposure to a given toxic agent or drug (ART exposure through PMTCT in our case) and an environmental or genetic change (Dettmer *et al.* 2007).

The overriding initial view of this third omics science is that it represents the final level in a biological system, where the metabolites act as functional entities, unlike their structural nucleic acid and protein precursors. Metabolites thus express functional processes in biological systems but it was soon realized (Raamsdonk *et al.* 2001) that the metabolite profiles also reflect the influence of the surrounding environment on the biological system. Metabolomics consequently provides new perspectives and a new

power to visualize metabolite information within the context of the physiological, developmental, pathological and survival status of a biological system.

One of the first papers to propose the potential of metabolomics in HIV and AIDS research came from the renowned laboratory of Gary Siuzdak (Department of Chemistry and Center for Mass Spectrometry, Scripps Research Institute, La Jolla, California, USA). The paper was an invited review on “*Biomarkers for NeuroAIDS: The Widening Scope of Metabolomics*” (Pendyala *et al.* 2007). In their paper the authors predicted that the new “omics” branch – metabolomics – “*may even become superior to any other postgenomics technologies for pattern-recognition analyses of biological samples.*”

Siuzdak and co-workers were apparently not aware of the first metabolomics study on HIV that was published in the year before their review. In a pioneering investigation on HIV infection, using ¹H-NMR (proton NMR) spectroscopy-based metabolomics and linear discriminant analysis, Prof Debra Meyer and co-workers from the University of Pretoria showed in their preliminary study on HIV that metabolomics data contain sufficient information to distinguish HIV-positive/AIDS (Hewer *et al.* 2006) patients on ART from uninfected individuals. These results were substantiated by a subsequent metabolomics investigation by the same group using 300- and 600-MHz NMR instruments and multinomial logistic regression classification of the metabolomics data (Philippeos *et al.* 2009). The chemical compounds corresponding to the sections of the NMR spectra that showed significant differences ($p < 0.05$) for infected and uninfected individuals were mainly lipids, including low-density lipoprotein (LDL) and very-low-density lipoprotein (VLDL). These observations concurred with the irregularities of lipodystrophy and hyperlipidemia common in HIV/AIDS patients on ART (Calza *et al.* 2003). The importance of the investigation of Meyer and co-workers was that it was the first study that gave analytical proof that metabolic changes induced by HIV and/or ART can be revealed by NMR metabolomics data.

A collaborative study between our own laboratory and that of Meyer and her co-workers, subsequently reported on a quantitative organic acid profile of HIV-infected individuals not receiving any ART (Williams *et al.* 2012). The study of Aurelia Williams was the first application of gas chromatography–mass spectrometry (GC–MS) metabolomics to analyse the organic acid profiles in urine of asymptomatic HIV-infected individuals ($n = 18$) compared to uninfected controls ($n = 21$). Several organic acids are well-established diagnostic biomarkers of mitochondrial dysfunction, which made the analysis of the organic acid metabolome well suited to monitoring the progressive disruption of mitochondrial structure and function during HIV infection.

Using a multifaceted analytical bioinformatics procedure, at least 10 of these metabolites could be linked to (1) disrupted mitochondrial metabolism, (2) changes in lipid metabolism and (3) oxidative stress, all of which are aberrations caused by HIV infection (Williams *et al.* 2012). Because of the role of the mitochondria in apoptosis, these authors also found that a higher level of this type of cell death in infected (compared to uninfected) individuals supported their GC–MS data (Williams *et al.* 2013). This early study thus demonstrated that mass spectrometry metabolomics detects biomarkers of mitochondrial dysfunction which could potentially be developed into indicators of HIV infection, perhaps also to monitor disease progression and the response to ART as envisaged for the present study.

A next important observation appeared in a review on metabolomics studies on HIV-infected patients. Based on the data covered in the review, *“it is not likely that a single metabolite (occurring in single or multiple pathways) will be detected as being responsible for HIV/AIDS induced metabolic change. Instead, the finding of synergistic molecules is more probable”* (Sitole *et al.* 2013). This observation concurred with an investigation on a metabolomic analysis of the cerebrospinal fluid on changes in the central nervous system of macaques infected with simian immunodeficiency virus (SIV) (Wikoff *et al.* 2008) and with the proposal that *multiple biomarkers* may be more advantageous than a single marker for biological information on disease mechanisms (Hollenbaugh *et al.* 2011). The investigation by Hollenbaugh and co-workers used liquid chromatography–tandem mass spectrometry (LC–MS/MS) technology to monitor metabolic alterations induced by HIV in macrophage-infected cells. HIV-infected CD4⁺ T cells showed higher glucose uptake and increases in several metabolite pool sizes, whereas HIV producing macrophages had substantial reductions in glucose uptake and lower steady-state glycolytic intermediates. These metabolic cell studies provided evidence that the two different HIV target cell types exhibit very different metabolic outcomes during viral production. Furthermore, epidemiological studies have shown that HIV-infected men on ART are at higher risk of incident diabetes mellitus compared with women with similar treatment histories. However, a metabolomics study to determine whether a gender difference in plasma amino acids, acyl-carnitines, and organic acids predictive of diabetes and impaired energy metabolism is present in HIV-infected persons on long-term ART, revealed that plasma metabolite levels did not significantly differ according to HIV status among the experimental subjects, suggesting the observed gender differences may not be specific to HIV infection (Koethe *et al.* 2016). It is noteworthy that the relationship between gender and metabolite levels thus appears to be biologically based rather

than HIV-specific, which sounds a note of caution regarding gender composition in selecting experimental subjects for metabolomics studies related to HIV or ART.

The risk involved in administering ART is one of the motivations for the present study. A few metabolomics studies involving ART support this concern. Ritonavir (RTV), an HIV protease inhibitor (PI), is currently used mainly as a booster of other PIs. However, indications are that RTV-boosted PI regimens may cause toxicity; it has been considered as a risk factor of HIV drug-induced liver injury (Bruno *et al.* 2006), as well as a cause of kidney damage (Shafi *et al.* 2011). The novelty of the RTV-metabolomics study was to explore the metabolic pathways of ritonavir that may be related to its toxicity. Ultra-performance liquid chromatography and time-of-flight mass spectrometry (UPLC-TOFMS) were used for metabolite analysis in the metabolomics study using mice as experimental animals. The metabolite fingerprint revealed 26 ritonavir-related metabolites of which half were new. Although the focus of the study was on drug metabolism, the presence of diverse and potential toxic RTV catabolites following PI treatment underscores the risks involved in ART.

A subsequent study focused on the biological consequences of ART (Munshi *et al.* 2013). While HIV primarily has an immune cell target, metabolic alterations with clinical consequences are also known to occur in HIV/AIDS patients. In this investigation, the researchers aimed to profile metabolite changes in plasma, urine and saliva of HIV/AIDS patients with and without ART. This was an NMR-based investigation using a Bruker Avance III spectrometer, operating at the field strength of 500 MHz, and thus closely resembles the metabolomics technology used in part of the present study. Principal components analysis (PCA) was performed using the first three principal components (PCs) on the plasma, urine and saliva data sets, which could distinguish between patients and controls on the basis of their NMR spectra. To describe maximum separation between predefined sample classes, the method of supervised partial least squares–discriminant analysis (PLS-DA) was applied. The PLS-DA was used to identify important discriminatory metabolites present in the experimental groups. Compared to healthy controls, the levels of sarcosine, methylmalonic acid, glucose, choline and aspartic acid were higher in the plasma of HIV/AIDS patients. Although the levels of these metabolites were reduced in patients receiving ART, they were still higher than those found in healthy subjects. Overall, 26 metabolites were differentially altered in any or two types of samples. Their results suggest that urinary neopterin as well as plasma choline and sarcosine could be used as metabolic biomarkers of HIV/AIDS infection. The researchers concluded that metabolites present in plasma, urine and to a lesser extent in saliva become differentially produced in

response to HIV infection, including significant changes in their levels following ART. Deregulated metabolic pathways related to the important metabolites were constructed, indicating that the perturbed metabolism may contribute to pathogenesis and disease outcome in HIV/AIDS patients.

A third metabolomics study on ART used ultra-high-performance liquid chromatography and tandem mass spectrometry (UHPLC/MS2/MS) as well as gas chromatography–mass spectrometry (GC–MS) for metabolite detection (Cassol *et al.* 2013). The aim of the study was to uncover biochemical details and underlying mechanisms which occur in HIV-infected individuals on ART. A complex 35-metabolite signature was observed which mapped lipid, amino acid, and nucleotide metabolism and distinguished from controls HIV patients in an advanced state of disease as well as being on PI-based ART. They further observed that many altered lipids, including bile acids, sulfated steroids, polyunsaturated fatty acids, and eicosanoids, were ligands of nuclear receptors that regulate metabolism and inflammation. Notwithstanding the complex perturbations observed as a consequence of HIV-ART, this study clearly indicated that untargeted metabolomics profiling represents a powerful tool for gaining a systems-level understanding of metabolic alterations. Cassol *et al.* (2013) thus concluded that the dysfunctional metabolism might involve organ systems that will be important for developing therapeutic strategies that target metabolic disease and chronic inflammation in HIV patients on ART.

To conclude this overview on metabolomics studies related to HIV and ART, I refer to a recent brief review by Williams *et al.* (2016) on alterations in the microbiome in HIV infection. They provide a comprehensive table listing the link between bacterial translocation and immune activation in HIV. Increased interest appears to have developed in changes in the microbiome due to HIV. As in metabolic studies, changes in the intestinal microbiome in HIV-positive subjects seemed to give somewhat inconsistent results. Currently, Williams *et al.* (2016) relate the inconsistency to small sample sizes, lack of appropriate controls or regional differences in dietary and environmental factors. Despite these limitations, this study concludes that the microbiome plays an important role in the pathogenesis of HIV disease, and that modulation of the microbiome presents an important therapeutic target to improve health in this patient population (Williams *et al.* 2016). It can be anticipated that metabolomics studies on the host–pathogen–gut interaction may provide a fruitful field for research and for novel insights into the complex profile of the HIV-ART scenario.

2.7 Aim and research question for the investigation

The above overview clearly supports Siuzdak's prediction that metabolomics may even become superior to any other post-genomics technologies for pattern-recognition analyses of biological samples (Pendyala *et al.* 2007), and confirms the richness in metabolomics approaches with regard to choice of analytical tools, experimental design and generation of novel information. I have already indicated that the introduction of ART to pregnant women to address PMTCT of HIV has become a cornerstone in public health and one of its greatest successes (Chasela *et al.* 2010). Despite the impressive reduction in HIV infection to infants following PMTCT, the treatment remains a matter of concern in view of the diverse aberrations reported for some neonates exposed to HIV-ART (Newell & Bunders 2013). Until now, no investigation has reported on the consequences of *in utero* exposure to HIV ART in newborns, either through conventional clinical biochemical or metabolomics studies. This leaves an important void in our understanding of HIV-ART exposure to infants, and is an aspect to be considered seriously in formulation of health policies to combat the spread of HIV.

The aim of the present investigation is therefore to address the concern following treatment for the purpose of PMTCT by exploring metabolic profiles in exposed neonates through a metabolomics approach.

The specific research question posed in this investigation is how to determine whether any HIV-ART-induced metabolic perturbations are discernible during birth, when the fetus is exposed to the trauma of birth hypoglycemia, keeping in mind that this is the very first start of a longitudinal study on the long-term effects of HIV-ART exposure.

The experimental approach followed to explore the research question applied metabolomics technology according to the following strategy: The first part of the investigation used untargeted ¹H-NMR. It was envisaged that the untargeted approach would disclose major metabolic perturbations due to *in utero* exposure to HIV-ART, if any, but might also indicate specific perturbations in some part of the metabolome. Any observation on a potential metabolic perturbation from the NMR study would provide a directive for a semi-targeted metabolomics approach to be addressed in the second part of the investigation. The final part of the study provides an integration of the insights into the consequences of HIV-ART as gained through the two metabolomics approaches.

REFERENCES

- Afran, L., Garcia Knight, M., Nduati, E., Urban, B. C., Heyderman, R. S., & Rowland-Jones, S. L. (2014). HIV-exposed uninfected children: a growing population with a vulnerable immune system?. *Clinical & Experimental Immunology*, 176(1), 11–22.
- Ahmed, S., Kim, M. H., & Abrams, E. J. (2013). Risks and benefits of lifelong antiretroviral treatment for pregnant and breastfeeding women: a review of the evidence for the option B+ approach. *Current Opinion in HIV and AIDS*, 8(5), 474–489.
- Alfadda, A. A., & Sallam, R. M. (2012). Reactive oxygen species in health and disease. *BioMed Research International*, 2012, 1–14.
- Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J. & Dauguet, C. (1983). Isolation of a T-Lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*. 220, 868-871.
- Beck, J. M., Rosen, M. J., & Peavy, H. H. (2001). Pulmonary complications of HIV infection: report of the fourth NHLBI workshop. *American Journal of Respiratory and Critical Care Medicine*, 164(11), 2120–2126.
- Bertolli, J., St Louis, M. E., Simonds, R. J., Nieburg, P., Kamenga, M., Brown, C., et al. (1996). Estimating the timing of mother-to-child transmission of human immunodeficiency virus in a breast-feeding population in Kinshasa, Zaire. *The Journal of Infectious Diseases*, 174, 722–726.
- Blanche, S., Tardieu, M., Benhammou, V., Warszawski, J., & Rustin, P. (2006). Mitochondrial dysfunction following perinatal exposure to nucleoside analogues. *AIDS*, 20(13), 1685–1690.
- Blanche, S., Tardieu, M., Rustin, P., Slama, A., Barret, B., Firtion, G., et al. (1999). Persistent mitochondrial dysfunction and perinatal exposure to antiretroviral nucleoside analogues. *Lancet*, 354, 1084–1089.
- Brinkman, K., Smeitink, J. A., Romijn, J. A., & Reiss, P. (1999). Mitochondrial toxicity induced by nucleoside-analogue reverse-transcriptase inhibitors is a key factor in the pathogenesis of antiretroviral-therapy-related lipodystrophy. *The Lancet*, 354(9184), 1112–1115.

- Brinkman, K., ter Hofstede, H. J., Burger, D. M., Smeitink, J. A., & Koopmans, P. P. (1998). Adverse effects of reverse transcriptase inhibitors: mitochondrial toxicity as common pathway. *AIDS*, 12(14), 1735–1744.
- Bruno, R., Sacchi, P., Maiocchi, L., Patruno, S., & Filice, G. (2006). Hepatotoxicity and antiretroviral therapy with protease inhibitors: A review. *Digestive and Liver Disease*, 38(6), 363–373.
- Bryson, Y. J., Luzuriaga, K., Sullivan, J. L., & Wara, D. W. (1992). Proposed definitions for in utero versus intrapartum transmission of HIV-1. *New England Journal of Medicine*, 327(17), 1246–1247.
- Burns, D. N., Landesman, S., Wright, D. J., Waters, D., Mitchell, R. M., Rubinstein, A., et al. (1997). Influence of other maternal variables on the relationship between maternal virus load and mother-to-infant transmission of human immunodeficiency virus type 1. *Journal of Infectious Diseases*, 175(5), 1206–1210.
- Calza, L., Manfredi, R., & Chiodo, F. (2003). Hyperlipidaemia in patients with HIV-1 infection receiving highly active antiretroviral therapy: epidemiology, pathogenesis, clinical course and management. *International journal of antimicrobial agents*, 22, 89–99.
- Carraro, S., Giordano, G., Reniero, F., Perilongo, G., & Baraldi, E. (2009). Metabolomics: a new frontier for research in pediatrics. *The Journal of Pediatrics*, 154, 638–644.
- Cassol, E., Misra, V., Holman, A., Kamat, A., Morgello, S., & Gabuzda, D. (2013). Plasma metabolomics identifies lipid abnormalities linked to markers of inflammation, microbial translocation, and hepatic function in HIV patients receiving protease inhibitors. *BMC Infectious Diseases*, 13(1), 1–27.
- CDC. 1981. Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men—New York City and California. *MMWR. Morbidity and mortality weekly report*, 30(25), 305–308.
- Chan, S. S., Santos, J. H., Meyer, J. N., Mandavilli, B. S., Cook, D. L., McCash, C. L., et al. (2007). Mitochondrial toxicity in hearts of CD-1 mice following perinatal exposure to AZT, 3TC, or AZT/3TC in combination. *Environmental and Molecular Mutagenesis*, 48(3-4), 190.

Chasela, C. S., Hudgens, M. G., Jamieson, D. J., Kayira, D., Hosseinipour, M. C., Kourtis, A. P., et al. (2010). Maternal or infant antiretroviral drugs to reduce HIV-1 transmission. *The New England Journal of Medicine*, 362(24), 2271–2281.

Chigwedere, P., Seage, G. R., Gruskin, S., Lee, T-H., & Essex, M. (2008). Estimating the lost benefits of antiretroviral drug used in South Africa. *Journal of Acquired Immune Deficiency Syndromes*, 00(0), 1–6.

Chirwa, W. C. (1998). Aliens and AIDS in Southern Africa: the Malawi-South Africa debate. *African Affairs*, 97(386), 53–79.

Clark, N. L., & Worger, W. H. (2013). *South Africa: the rise and fall of apartheid*. Routledge.

Cohen, M. S., Shaw, G. M., McMichael, A. J., & Haynes, B. F. (2011). Acute HIV-1 infection. *New England Journal of Medicine*, 364(20), 1943–1954.

Colbers, A. P., Hawkins, D. A., Gengelmaier, A., Kabeya, K., Rockstroh, J. K., Wyen, C., et al. (2013). The pharmacokinetics, safety and efficacy of tenofovir and emtricitabine in HIV-1-infected pregnant women. *AIDS*, 27(5), 739–748.

Connor, E. M., Sperling, R. S., Gelber, R., Kiselev, P., Scott, G., O'Sullivan, M. J., et al. (1994). Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. *New England Journal of Medicine*, 331(18), 1173–1180.

Cooper, E. R., Charurat, M., Mofenson, L., Hanson, I. C., Pitt, J., Diaz, C., et al. (2002). Combination antiretroviral strategies for the treatment of pregnant HIV-1-infected women and prevention of perinatal HIV-1 transmission. *Journal Acquired Immune Deficiency Syndrome*, 29, 484–494.

Coovadia, H., Jewkes, R., Barron, P., Sanders, D., & McIntyre, D. (2009). The health and health systems of South Africa: historical roots of current public health challenges. *Lancet*, 374, 817–834.

Crush, J., Williams, V. & Peberdy, S. (2005). Migration in southern Africa. *Policy Analysis and Research Programme of the Global Commission on International Migration*. Available at: <http://www.gcim.org> (Accessed: 15th February 2015).

- Dagan, T., Sable, C., Bray, J., & Gerschenson, M. (2002). Mitochondrial dysfunction and antiretroviral nucleoside analog toxicities: what is the evidence?. *Mitochondrion*, 1(5), 397–412.
- De Cock, K. M., Fowler, M. G., Mercier, E., de Vincenzi, I., Saba, J., Hoff, E., et al. (2000). Prevention of mother-to-child HIV transmission in resource-poor countries: translating research into policy and practice. *JAMA*, 283(9), 1175–1182.
- De Santis, M., Carducci, B., De Santis, L., Cavaliere, A. F., & Straface, G. (2002). Periconceptional exposure to efavirenz and neural tube defects. *Archives of Internal Medicine*, 162(3), 355–356.
- Dettmer, K., Aronov, P. A., & Hammock, B. D. (2007). Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26(1), 51–78.
- Doherty, M., Ford, N., Vitoria, M., Weiler, G., & Hirschall, G. (2013). The 2013 WHO guidelines for antiretroviral therapy: evidence-based recommendations to face new epidemic realities. *Current Opinion HIV AIDS*, 8(6), 528–534.
- Drake, A. L., Wagner, A., Richardson, B., & John-Stewart, G. (2014). Incident HIV during pregnancy and postpartum and risk of mother-to-child HIV transmission: a systemic review and meta-analysis. *PLoS One*, 11, 1–16.
- Drewes, J. L., Meulendyke, K. A., Liao, Z., Witwer, K. W., Gama, L., Ubaida-Mohien, C., et al. (2015). Quinolinic acid/tryptophan ratios predict neurological disease in SIV-infected macaques and remain elevated in the brain under cART. *Journal of Neurovirology*, 21(4), 449–463.
- Estrada, V., & Portilla, J. (2010). Dyslipidemia related to antiretroviral therapy. *AIDS Reviews*, 13(1), 49–56.
- Fauci, A. S. (2003). HIV and AIDS: 20 years of science. *Nature Medicine*, 9(7), 839–843.
- Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. *Plant Molecular Biology*, 48(1-2), 155–171.
- Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F., et al. (1999). Origin of HIV-1 in the chimpanzee *Pan troglodytes*. *Nature*, 397, 436–441.

- Gao, F., Yue, L., Robertson, D. L., Hill, S. C., Hui, H., Biggar, R. J., et al. (1994). Genetic diversity of human immunodeficiency virus type 2: evidence for distinct sequence subtypes with differences in virus biology. *Journal of Virology*, 68(11), 7433–7447.
- Giaquinto, C., De Romeo, A., Giacomet, V., Rampon, O., Ruga, E., Burlina, A., et al. (2001). Lactic acid levels in children perinatally treated with antiretroviral agents to prevent HIV transmission. *AIDS*, 15(8), 1074–1075.
- Gostner, J. M., Becker, K., Kurz, K., & Fuchs, D. (2015). Disturbed amino acid metabolism in HIV: association with neuropsychiatric symptoms. *Frontiers in Psychiatry*, 6(97), 1–10.
- Greaves, L. C., Reeve, A. K., Taylor, R. W., & Turnbull, D. M. (2012). Mitochondrial DNA and disease. *The Journal of Pathology*, 226(2), 274–286.
- Grunfeld, C., Pang, M. I. Y. I. N., Doerrler, W., Shigenaga, J. K., Jensen, P., & Feingold, K. R. (1992). Lipids, lipoproteins, triglyceride clearance, and cytokines in human immunodeficiency virus infection and the acquired immunodeficiency syndrome. *The Journal of Clinical Endocrinology & Metabolism*, 74(5), 1045–1052.
- Hahn, B. H., Shaw, G. M., De, K. M., & Sharp, P. M. (2000). AIDS as a zoonosis: scientific and public health implications. *Science*, 287, 607–614.
- Hammett, T. M., & Bronson, R. T. (2016). Unrecognized “AIDS” in monkeys, 1969–1980: explanations and implications. *American Journal of Public Health*, 106(6), 1015–1022.
- Hewer, R., Vorster, J., Steffens, F. E., & Meyer, D. (2006). Applying biofluid 1 H NMR-based metabonomic techniques to distinguish between HIV-1 positive/AIDS patients on antiretroviral treatment and HIV-1 negative individuals. *Journal of Pharmaceutical and Biomedical Analysis*, 41, 1442–1446.
- Hirsch, V. M., Olmsted, R. A., Murphey-Corb, M., Purcell, R. H., & Johnson, P. R. (1989). An African primate lentivirus (SIVsm) closely related to HIV-2.
- Hollenbaugh, J. A., Munger, J., & Kim, B. (2011). Metabolite profiles of human immunodeficiency virus infected CD4+ T cells and macrophages using LC–MS/MS analysis. *Virology*, 415(2), 153–159.

- Ivanov, A. V., Valuev-Elliston, V. T., Ivanova, O. N., Kochetkov, S. N., Starodubova, E. S., Bartosch, B., et al. (2016). Oxidative stress during HIV infection: mechanisms and consequences. *Oxidative Medicine and Cellular Longevity*, 2016, 1–18.
- Jourdain, G., Mary, J. Y., Le Coeur, S., Ngo-Giang-Huong, N., Yuthavisuthi, P., Limtrakul, A., et al. (2007). Risk factors for in utero or intrapartum mother-to-child transmission of human immunodeficiency virus type 1 in Thailand. *Journal of Infectious Diseases*, 96(11), 1629–1636.
- Karim, Q. A., & Karim, S. S. A. (2002). The evolving HIV epidemic in South Africa. *International Journal of Epidemiology*, 31(1), 37–40.
- Kaul, M., Garden, G. A., & Lipton, S. A. (2001). Pathways to neuronal injury and apoptosis in HIV-associated dementia. *Nature*, 410, 988–994.
- Kellerman, S. E., Ahmed, S., Feeley-Summerl, T., Jay, J., Kim, M. H., Koumans, E., et al. (2013). Beyond PMTCT: keeping HIV exposed and positive children healthy and alive. *AIDS*, 27(0 2), S225.
- Kirmse, B., Baumgart, S., & Rakhmanina, N. (2013b). Metabolic and mitochondrial effects of antiretroviral drug exposure in pregnancy and postpartum: implications for fetal and future health. In *Seminars in Fetal and Neonatal Medicine*, 18(1), 48–55.
- Kirmse, B., Hobbs, C. V., Peter, I., LaPlante, B., Caggana, M., Kloke, K., et al. (2013a). Abnormal newborn screening data and acylcarnitines in HIV/ARV-exposed infants. *The Pediatric Infectious Disease Journal*, 32(2), 146–150.
- Koethe, J. R., Jenkins, C. A., Petucci, C., Culver, J., Shepherd, B. E., & Sterling, T. R. (2016). Superior glucose tolerance and metabolomic profiles, independent of adiposity, in HIV-infected women compared with men on antiretroviral therapy. *Medicine*, 95(19), e3634.
- Kourtis, A. P., Lee, F. K., Abrams, E. J., Jamieson, D. J., & Bulterys, M. (2006). Mother-to-child transmission of HIV-1: timing and implications for prevention. *The Lancet Infectious Diseases*, 6(11), 726–732.
- Lewis, W., Day, B. J., & Copeland, W. C. (2003). Mitochondrial toxicity of NRTI antiviral drugs: an integrated cellular perspective. *Nature Reviews Drug Discovery*, 2(10), 812–822.

Lim, S. E., & Copeland, W. C. (2001). Differential incorporation and removal of antiviral deoxynucleotides by human DNA polymerase γ . *Journal of Biological Chemistry*, 276(26), 23616–23623.

Lurie, M. N., Williams, B. G., Zuma, K., Mkaya-Mwamburi, D., Garnett, G. P., Sweat, M. D., et al. (2003). Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa. *AIDS*, 17(15), 2245–2252.

Mandas, A., Iorio, E. L., Congiu, M. G., Balestrieri, C., Mereu, A., Cau, D., et al. (2009). Oxidative imbalance in HIV-1 infected patients treated with antiretroviral therapy. *Journal of Biomedicine and Biotechnology*, 2009, 1–7.

Mas, C. M., Miller, T. L., Cordero, C., Dauphin, D., White, M. B., Vila, C. K., et al. (2012). The effects of fetal and childhood exposure to antiretroviral agents. *Journal of AIDS & Clinical Research*, (Supplement 2), 2–10.

McComsey, G. A., Ward, D. J., Hesselthaler, S. M., Sension, M. G., Shalit, P., Lonergan, J. T., et al. (2004). Improvement in lipoatrophy associated with highly active antiretroviral therapy in human immunodeficiency virus—infected patients switched from stavudine to abacavir or zidovudine: The results of the TARHEEL study. *Clinical Infectious Diseases*, 38(2), 263–270.

Mofenson, L. M., Lambert, J. S., Stiehmel, E. R., Bethel, J., Meyer, W. A., Whitehouse, J., et al. (1999). Risk factors for perinatal transmission of human immunodeficiency virus type 1 in women treated with zidovudine. *New England Journal of Medicine*, 341(6), 385–393.

Moodley, D., Esterhuizen, T. M., Pather, T. Chetty, V., & Ngaleka, L. (2009). High HIV incidence during pregnancy: compelling reason for repeat HIV testing. *AIDS*, 23, 1255–1259.

Mudd, L. M., Holzman, C. B., Catov, J. M., Senagore, P. K., & Evans, R. W. (2012). Maternal lipids at mid-pregnancy and the risk of preterm delivery. *Acta Obstetrica et Gynecologica Scandinavica*, 91(6), 726–735.

Mulligan, K., Grunfeld, C., Tai, V. W., Algren, H., Pang, M., Chernoff, D. N., et al. (2000). Hyperlipidemia and insulin resistance are induced by protease inhibitors independent of changes in body composition in patients with HIV infection. *Journal of Acquired Immune Deficiency Syndromes*, 23(1), 35–43.

- Munshi, S. U., Rewari, B. B., Bhavesh, N. S., & Jameel, S. (2013). Nuclear magnetic resonance based profiling of biofluids reveals metabolic dysregulation in HIV-infected persons and those on anti-retroviral therapy. *PloS One*, 8(5), 1–9.
- Nattrass, N. (2008). AIDS and the Scientific governance of medicine in post-apartheid South Africa. *African Affairs*, 107, 157–176.
- Newell, M. L. (2001). Prevention of mother-to-child transmission of HIV: challenges for the current decade. *Bulletin of the World Health Organization*, 79(12), 1138–1144.
- Newell, M. L., & Bunders, M. J. (2013). Safety of antiretroviral drugs in pregnancy and breastfeeding for mother and child. *Current Opinion in HIV and AIDS*, 8(5), 504–510.
- Nicholson, J. K., Lindon, J. C., & Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29(11), 1181–1189.
- Noguera, A., Fortuny, C., Muñoz-Almagro, C., Sanchez, E., Vilaseca, M. A., Artuch, R., et al. (2004). Hyperlactatemia in human immunodeficiency virus–uninfected infants who are exposed to antiretrovirals. *Pediatrics*, 114(5), e598–e603.
- Oliver, S. G., Winson, M. K., Kell, D. B., & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, 16, 373–378.
- Pendyala, G., Want, E. J., Webb, W., Siuzdak, G., & Fox, H. S. (2007). Biomarkers for neuroAIDS: the widening scope of metabolomics. *Journal of Neuroimmune Pharmacology*, 2(1), 72–80.
- Petropoulou, H., Stratigos, A. J., & Katsambas, A. D. (2006). Human immunodeficiency virus infection and pregnancy. *Clinics in Dermatology*, 24(6), 536–542.
- Philippeos, C., Steffens, F. E., & Meyer, D. (2009). Comparative ¹H NMR-based metabonomic analysis of HIV-1 sera. *Journal of Biomolecular NMR*, 44, 127–137.
- Pliner, V., Weedon, J., Thomas, P. A., Steketee, R. W., Abrams, E. J., Lambert, G., et al. (1998). Incubation period of HIV-1 in perinatally infected children. *AIDS*, 12(7), 759–766.

Poirier, M. C., Divi, R. L., Al-Harathi, L., Olivero, O. A., Nguyen, V., Walker, B., et al. (2003). Long-term mitochondrial toxicity in HIV-uninfected infants born to HIV-infected mothers. *Journal of Acquired Immune Deficiency Syndromes*, 33(2), 175–183.

Poirier, M.C., Gibbons, A. T., Rugeles, M. T., Andere-Schmutz, I., & Blanche, S. (2015). Fetal consequences of maternal antiretroviral nucleoside reverse transcriptase inhibitor use in human and nonhuman primate pregnancy. *Current Opinion in Pediatrics*, 27(2), 233–239.

Popovic, M., Sarngadharan, M. G., Read, E., & Gallo, R. C. (1984). Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science*, 224, 497–500.

Raamsdonk, L. M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M. C., et al. (2001). A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology*, 19(1), 45–50.

Read, J. S., Scherpbier, H. J., & Boer, K. (1999). The mode of delivery and the risk of vertical transmission of human immunodeficiency virus type 1: a meta-analysis of 15 prospective cohort studies: the International Perinatal HIV Group. *The New England Journal of Medicine*, 340, 977–987.

Saitoh, A., Hull, A. D., Franklin, P., & Spector, S. A. (2005). Myelomeningocele in an infant with intrauterine exposure to efavirenz. *Journal of Perinatology*, 25(8), 555–556.

Scalfaro, P., Chesaux, J. J., Buchwalder, P. A., Biollaz, J., & Micheli, J. L. (1998). Severe transient neonatal lactic acidosis during prophylactic zidovudine treatment. *Intensive Care Medicine*, 24(3), 247–250.

Shafi, T., Choi, M. J., Racusen, L. C., Spacek, L. A., Berry, C., Atta, M., & Fine, D. M. (2011). Ritonavir-induced acute kidney injury: kidney biopsy findings and review of literature. *Clinical Nephrology*, 75(1), 60–64.

Shah, A., Kumar, S., Simon, S. D., Singh, D. P., & Kumar, A. (2013). HIV gp120-and methamphetamine-mediated oxidative stress induces astrocyte apoptosis via cytochrome P450 2E1. *Cell Death & Disease*, 4(10), e850.

Sharp, P. M., & Hahn, B. H. (2011). Origins of HIV and the AIDS pandemic. *Cold Spring Harbor Perspectives in Medicine*, 1(1), 1–22.

Sharp, P. M., Bailes, E., Chaudhuri, R. R., Rodenburg, C. M., Santiago, M. O., & Hahn, B. H. (2001). The origins of acquired immune deficiency syndrome viruses: where and when? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 356(1410), 867–876.

Simelela, N. P., & Venter, W. D. F. (2014). A brief history of South Africa's response to AIDS. *South African Medical Journal*, 104(3), 249–251.

Simon, V., Ho, D. D., & Karim, Q. A. (2006). HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. *The Lancet*, 368(9534), 489–504.

Simonon, A., Lepage, P., Karita, E., Hitimana, D. G., Dabis, F., Msellati, P., et al. (1994). An assessment of the timing of mother-to-child transmission of human immunodeficiency virus type 1 by means of polymerase chain reaction. *Journal of Acquired Immune Deficiency Syndromes*, 7(9), 952–957.

Sitole, L. J., Williams, A. A., & Meyer, D. (2013). Metabonomic analysis of HIV-infected biofluids. *Molecular BioSystems*, 9(1), 18–28.

Slogrove, A. L., Goetghebuer, T., Cotton, M. F., Singe, J., & Bettinger, J. A. (2016). Pattern of infectious morbidity in HIV-exposed uninfected infants and children. *Frontiers in Immunology*, 7(164), 1-8.

Sperling, R. S., Shapiro, D. E., Coombs, R. W., Todd, J. A., Herman, S. A., McSherry, G. D., et al. (1996). Maternal viral load, zidovudine treatment, and the risk of transmission of human immunodeficiency virus type 1 from mother to infant. *New England Journal of Medicine*, 335(22), 1621–1629.

Sperling, R. S., Shapiro, D. E., McSherry, G. D., Britto, P., Cunningham, B. E., Culnane, M., et al. (1998). Safety of the maternal–infant zidovudine regimen utilized in the Pediatric AIDS Clinical Trial Group 076 Study. *AIDS*, 12(14), 1805–1813.

Sripan, P., Le Coeur, S., Amzal, B., Ingsrisawang, L., Traisathit, P., Ngo-Giang-Huong, N., et al. (2015). Modeling of in-utero and intra-partum transmissions to evaluate the efficacy of interventions for the prevention of perinatal HIV. *PLoS One*, 10(5), 1–16.

Thorne, C., & Newell, M. L. (2007). Safety of agents used to prevent mother-to-child transmission of HIV. *Drug Safety*, 30(3), 203–213.

Tolle, M. A., & Dewey, D. (2010). Prevention of mother-to-child transmission of HIV infection. *HIV Curriculum for the Health Professional, Baylor International Pediatric AIDS Initiative, Houston*, 90-119.

Tsague, L., & Abrams, E. J. (2014). Commentary: Antiretroviral treatment for pregnant and breastfeeding women—the shifting paradigm. *AIDS*, 28(supplement 2), S119-S121.

UNAIDS. (2011) Countdown to zero: Global plan towards the elimination of new HIV infections among children by 2015 and keeping their mothers alive. URL: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/20110609_jc2137_global-plan-elimination-hiv-children_en.pdf. (Accessed September 2016).

UNAIDS. (2015). How AIDS changed everything. *Geneva, Switzerland: UNAIDS*. URL: www.unaids.org/sites/default/files/media_asset/MDG6Report_en.pdf. (Accessed September 2016).

Vilaseca, M. A., Artuch, R., Sierra, C., Pineda, J., Lopez-Vilches, M. A., Munoz-Almagro, C., & Fortuny, C. (2003). Low serum carnitine in HIV-infected children on antiretroviral treatment. *European Journal of Clinical Nutrition*, 57(10), 1317–1322.

Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S., & Alizon, M. (1985). Nucleotide sequence of the AIDS virus, LAV. *Cell*, 40(1), 9–17.

WHO. (2010). Antiretroviral drugs for treating pregnant women and preventing HIV infection in infants: recommendations for a public health approach-2010 version. World Health Organization. <http://www.who.int/hiv/pub/mtct/antiretroviral2010/en/index.html>. (Accessed September 2016)

WHO. (2010). PMTCT strategic vision 2010–2015: preventing mother-to-child transmission of HIV to reach the UNGASS and Millennium Development Goals.

WHO. (2015). Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. September 2015.

Wikoff, W. R., Gangoiti, J. A., Barshop, B. A., & Siuzdak, G. (2007). Metabolomics identifies perturbations in human disorders of propionate metabolism. *Clinical Chemistry*, 53(12), 2169–2176.

Wikoff, W. R., Pendyala, G., Siuzdak, G., & Fox, H. S. (2008). Metabolomic analysis of the cerebrospinal fluid reveals changes in phospholipase expression in the CNS of SIV-infected macaques. *The Journal of Clinical Investigation*, 118(7), 2661–2669.

Wilkinson, E., Engelbrecht, S., & de Oliveira, T. (2015). History and origin of the HIV-1 subtype C epidemic in South Africa and the greater southern African region. *Scientific Reports*, 5.

Williams, A. A. (2012). *Metabonomics profile and corresponding immune parameters of HIV infected individuals* (Doctoral dissertation, University of Pretoria).

Williams, A., Koekemoer, G., Lindeque, Z., Reinecke, C., & Meyer, D. (2012). Qualitative serum organic acid profiles of HIV-infected individuals not on antiretroviral treatment. *Metabolomics*, 8(5), 804–818.

Williams, A., Steffens, F., Reinecke, C., & Meyer, D. (2013). The Th1/Th2/Th17 cytokine profile of HIV-infected individuals: a multivariate cytokinomics approach. *Cytokine*, 61(2), 521–526.

Williams, B., Landay, A., & Presti, R. M. (2016). Microbiome alterations in HIV infection a review. *Cellular Microbiology*, 18(5), 645–651.

Zangerle, R., Kurz, K., Neurauter, G., Kitchen, M., Sarcletti, M., & Fuchs, D. (2010). Increased blood phenylalanine to tyrosine ratio in HIV-1 infection and correction following effective antiretroviral therapy. *Brain, Behavior, and Immunity*, 24(3), 403–408.

CHAPTER 3 METABOLOMICS METHODOLOGY

3.1 Outline

Metabolomic investigations analyse the state of biological systems by monitoring the metabolome. This can be implemented through the use of the following approaches: (1) targeted metabolomics (which involves quantitative measurement of a limited number of compounds); (2) metabolite profiling (based on non-targeted, semi-quantitative measurement of major metabolites, recordable above the detection limit); and (3) metabolic fingerprinting (non-targeted pattern recognition without metabolite identification). Shulaev (2006) includes these aspects extensively in his review, which covers major approaches in metabolomics, an overview of the modern analytical tools used for data generation, as well as the important context of bioinformatics needs. In this chapter I focus on the use of metabolomics as an analytical methodology and its use in the biological sciences, with occasional reference to the context of my own study.

3.2 Metabolomics: from small molecules to data to big ideas

Metabolomics (Fiehn 2002), the measurement of changing metabolite levels in cell systems from biological samples, is evolving from being an emerging science to becoming well-established. As Aristotle once observed, *“the whole is more than the sum of its parts”* (Aristotle ca. 350 BCE) – metabolomics has transformed the reductionist approach of identifying and recording changes of a few or a single metabolite linked to a particular pathway, to a holistic approach by measuring many detectable metabolites at the same time (Kell & Oliver 2004). Griffin *et al.* (2006) informs us of how the terms metabolome and metabolomics (Oliver *et al.* 1998) – and the related term metabonomics (Nicholson *et al.* 1999) – were coined at the end of the 1990s, *“to describe the development of approaches which aim to measure all the metabolites that are present within a cell, tissue or organism during a genetic modification or physiological stimulus”*. A review by Gates and Sweeley in 1978, concerning qualitative metabolic profiling, however, indicates that quantitative *“metabolic profiling”* had already become practised in the late 1940s and early 1950s; the analysis of the metabolome – the total collection of metabolites in a cell, tissue or organism – was already well-accomplished at that time. Since then the defined field of metabolomics has expanded dramatically and has continued to evolve through contributions from various disciplines, such as biology, analytical chemistry and

statistics. As shown in Figure 1, Robert Powers (2015), in an editorial preface in *Current Metabolomics*, indicates the growth of metabolomic publications through a PubMed search using the terms “*Metabolomics*” and “*Metabonomics*”. Similarly, Kell and Oliver (figure not shown), in their 2016 review, disclose that “*The annual numbers of papers with the term metabolom in their title or abstract continue to rise, and in 2015 amounted at Web of Knowledge to 3130*” (Kell & Oliver 2016).

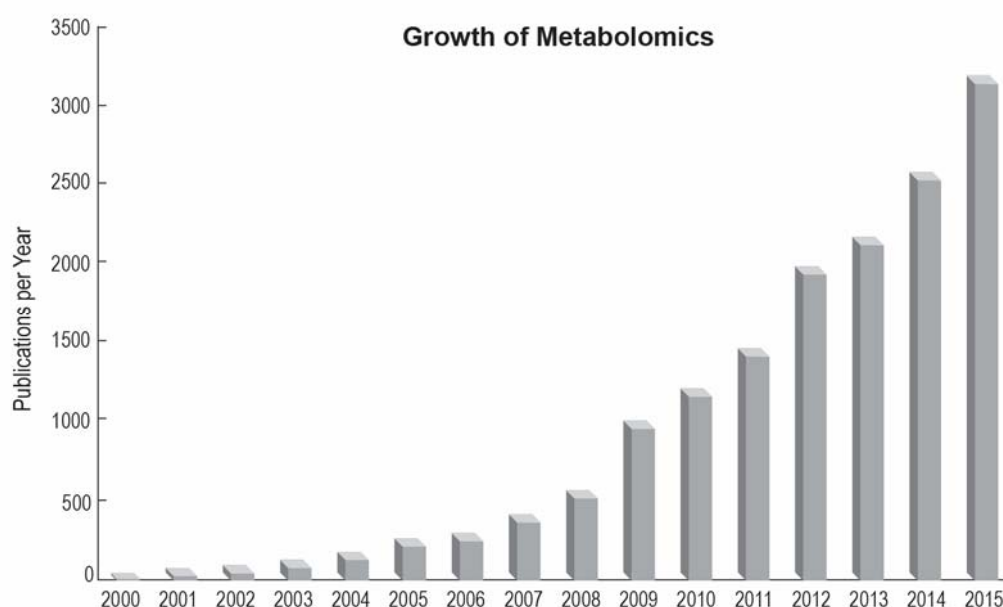


Figure 3–1: Number of manuscripts returned from a PubMed search using the terms ‘Metabolomics’ and ‘Metabonomics’ [Adapted from Powers, (2002- 2015) and Kell & Oliver 2014-2015)].

The ability of metabolomics to profile low-molecular-weight metabolites quantitatively in complex biological systems of all organisms has increased in popularity as the preferred analytical method for the investigation of metabolic perturbations (Moco *et al.* 2013). Plant biologists, more than any other biological research community, have rapidly adopted metabolomics as an effective study tool (Fiehn 2002). This is because plants in general have relatively large and complex metabolomes (Bino *et al.* 2004).

For decades in research medicine, one of the greatest obstacles faced was the difficulty to comprehend the relationship and connections between genetic variations and environmental triggers of disease (Nicholson 2006). The approaches of systems biology have been able to overcome this challenge by unfolding logical frameworks which are able to elucidate disease aetiology and reveal underlying networks between

seemingly dissimilar disease states (Loscalzo *et al.* 2007). With metabolomics regarded as a powerful top-down systems biology instrument, it is plausible to believe that a metabolic phenotype can offer information about an individual's metabolic state (in terms of the products of genetic and environmental conditions), that are unobtainable directly from the genome (DNA) or proteome (proteins) of an individual (Nicholson 2006). Jeremy Nicholson and co-workers developed a conceptual representation in a Forum Paper in the journal *Cell* (Holmes *et al.* 2008) to encapsulate the interactions between genes and the environment to indicate their influence on health and disease. Metabolic phenotypes (designated by Holmes *et al.* (2008) as “*metabotypes*”) can be defined by profiling small molecules (<1 kDa in size) in biofluids (usually urine or blood) by spectroscopic/spectrometric techniques. They indicate the way in which metabolic phenotypes are influenced by intrinsic factors (of genetic origin) and extrinsic factors (of environmental influences) that determine the health status/disease state/disease risk of an individual or group. Measuring and modelling the profile of all metabolites (the metabolome) in disease may provide insights for diagnosis, treatment and prognosis; this condition is also expressed as the *conditional metabolic phenotype*, which is a new concept introduced by Holmes. They define the conditional metabolic phenotype as “a characteristic metabolite profile reflecting the host genome and its interaction with environmental factors, diet, and the gut microbiome”. In my study the host genome is the genetic identity of the fetus, exposed to the intrauterine environment of the maternal HIV, followed by the effect of the application of ART, introduced in the second half of pregnancy (second trimester) which could potentially lead to epigenetic changes related to the in utero environmental exposure. I have used the conceptual model of Holmes *et al.* (2008) to develop a representation that depicts the main lines of thought and experimental development in my thesis, as shown in Figure 3-2.

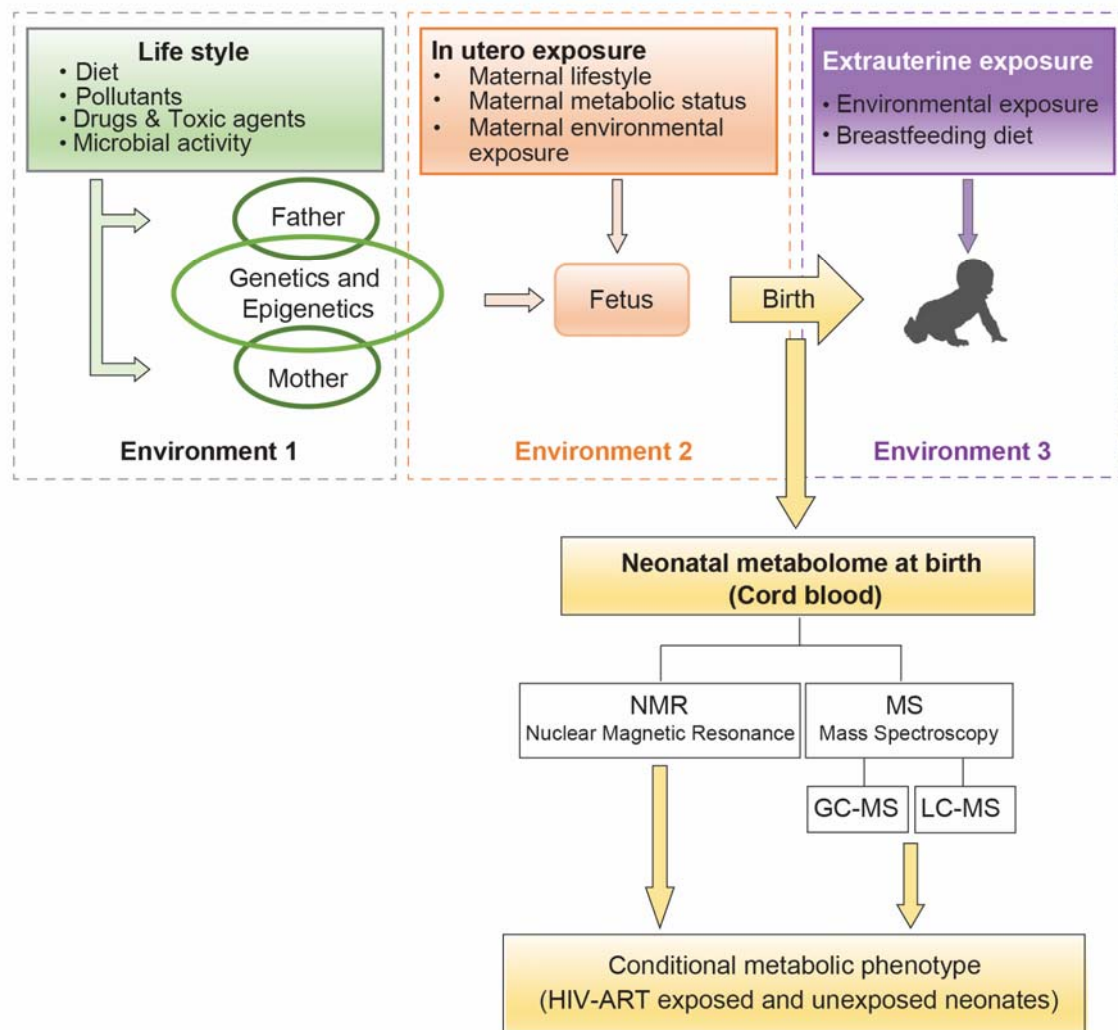


Figure 3–2: A conceptual representation of the evolution of the metabolic phenotype as a consequence of an *in utero* exposure (or no exposure) to maternal HIV-ART. Environment 1 relates to the genetic and epigenetic characteristics of the father and mother, transferred at conception to the fertilized ovum. Environment 2 indicates the prenatal environment of the fetus without exposure (controls) or exposed to HIV-positive mothers treated with ART. Environment 3 indicates the potential long-term consequences of HIV-ART on the infants, as they develop towards adulthood.

The genetic predisposition of the fetus results from specific genetic variations inherited from its parents (Figure 3-2, Environment 1). This predisposition may contribute to the susceptibility of the fetus to complexities within the intrauterine environment (Figure 3-2, Environment 2) whereas the predisposing genetic variation of another fetus, even

within the same family, will not be susceptible to a comparable intrauterine environment. Two decades ago, Barker stated that: "*The womb may be more important than the home*", and recommends that "research should be redirected towards the intrauterine environment rather than the environment in later childhood" (Barker 1990). A related view was expressed in a review by Gluckman and Hanson (2004) on the "*Effect of in utero and early-life conditions on adult health and disease*". They state that fetal adaptive responses to an adverse intrauterine environment, cues taken from the maternal health or physical state or perinatal responses such as changes in metabolism and hormone production, can lead to consequences of long-term disease risks later in adult life (Gluckman & Hanson 2004). It was formerly hypothesized by Barker and co-workers (2002) that environmental factors, particularly nutrition, act in early life to programme the risks for adverse health outcomes in adult life. This awareness of the role of nutrition and lifestyle in health and disease risk management has increased over the past decade with more emphasis being placed on the prevention of metabolic disorder (Holmes *et al.* 2008) (Figure 3-2, Environment 3). Metabolomic-driven clinical studies, which focus on biomarker discovery, are in a position to help us better understand genetic–environment–health interactions (Moco *et al.* 2013).

Against this background, the metabolomics study for this thesis was designed to discover if there is any effect on the homeostasis of neonates exposed *in utero* to HIV-ART or unexposed to the pathogen (HIV) and to possible drug toxicity (introduced via ART). An important aspect of this study is the collection of cord blood from the umbilical cord at birth, aimed at obtaining information on the conditional metabolic profile of the unexposed and exposed neonates at the critical point of birth.

3.3 Approaches to metabolomics

3.3.1 Background

Metabolomics investigations, like all good scientific experiments, require a meticulously well-thought-out experimental design (Goodacre *et al.* 2007). In a metabolomics investigation the main aspects of an experimental design are summarized in a generic "*workflow*", which typically includes sample preparation, the necessary analytical instruments capable of detecting and isolating metabolite compounds of interest as well as data processing and data analysis. The design of

various workflows in this thesis depended on the purpose of the subsection described, and are shown in Chapter 4 for the NMR analysis (Figure 4-1) and in Chapter 5 for the development of KEMREP (Figure 5-2). The biological question of the present study is new and unique, as it deals with a metabolomics study in the field of neonatology with a focus on the global health issues of PMTCT. Addressing the biological question required the participation of several experts in the design and implementation of the project. I therefore structured this review on the logic which underpins a metabolomics workflow, and summarize this logic in the design of my metabolomics study by means of the diagram shown in Figure 3-3, and discussed below following the corresponding sequential subsections.

3.4 Approaches to metabolomics

3.4.1 Define: Research question and design

Metabolomic studies which aim to describe the metabolome profile of any individual or population group in terms of factors such as their gender, age diet, gut microbiota, latent disease, stress levels and medication need to formulate a well-structured experimental design (the workflow). Analysts who partake in metabolomics investigations, which include the processing of samples, metabolite analysis and data processing, need to make sure that their choice of sample (biological fluid or other biological material) is representative of qualitative and quantitative variations of phenotypic molecular markers that will be obtained during metabolite extraction (Dunn *et al.* 2005; Holmes *et al.* 2008). The present study, being in the field of neonatology, posed a particular challenge in the selection of the experimental subjects from which the biofluids would be collected, including the procedures to be followed with sampling.

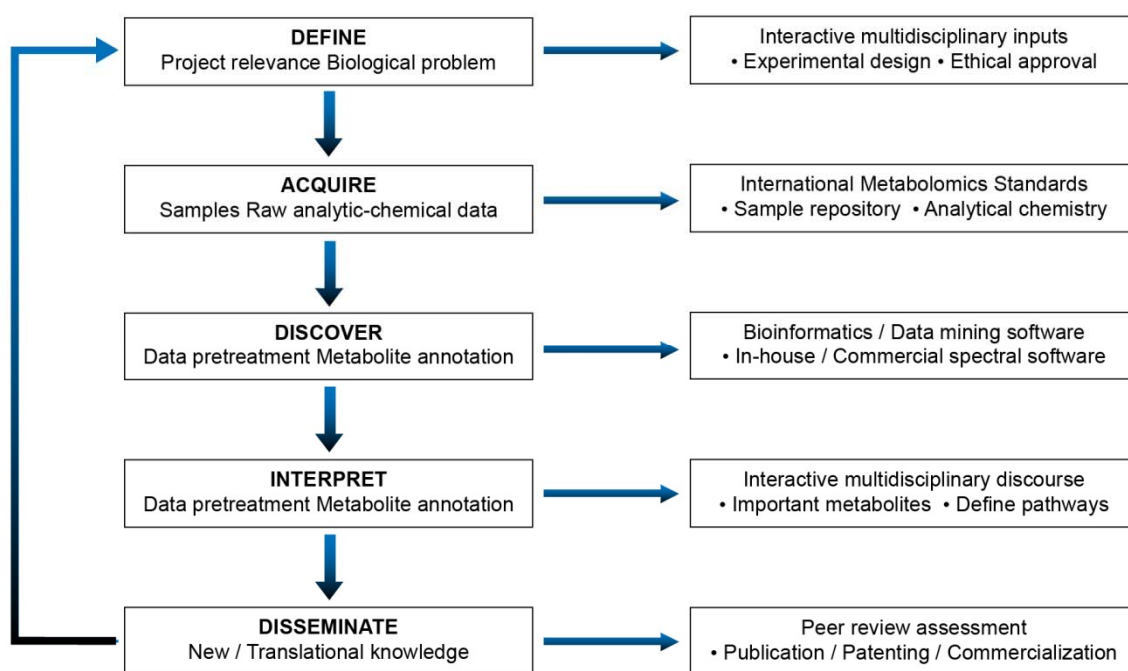


Figure 3–3: Schematic representation of the workflow for the metabolomics investigation. The design is presented as an overview of a pipeline of a metabolomics platform ranging from the design, performance, analysis and interpretational aspects, dissemination of metabolomics experiments (highlighted by rectangular boxes shown to the left of the figure) and their attendant requirements and multidisciplinary involvements (highlighted by the rectangular boxes shown to the right of the figure).

Vassilios Fanos from the Department of Neonatology, University of Cagliari, Italy, was one of the first paediatricians to recognize the potential of metabolomics in neonatology: “*The metabolomics approach, together with transcriptomics and proteomics, will have substantial impact on development of diagnostics, therapeutics and drug development and may be an important new tool in neonatology*” (Atzori *et al.* 2009). In neonatology, metabolomics has been applied in assessing responses to environmental stress (Hashimoto *et al.* 2013), the monitoring of different growth stages and homeostasis (Fanos *et al.* 2011), toxicology (Robertson *et al.* 2010), probe metabolic adaptations associated with premature birth, and maternal-fetal nutrient exchange (Tea *et al.* 2012). Also involved is the disease profiling of newborns with

asphyxia, and hypoxic ischaemic encephalopathy (Reinke *et al.* 2013; Walsh *et al.* 2012; Fanos *et al.* 2013).

The application of NMR in cord blood analysis is reviewed in Chapter 7 in conjunction with results from the present investigation (Chapter 4). However, Fanos noticed that it cannot be ignored that even though there has been an exponential growth in metabolomics as it has attracted significant attention, yet the use of metabolomics in perinatal and paediatric medicine is still an “embryonic” science (Fanos *et al.* 2013).

The top-left box in Figure 3-3 defines the experimental relevance and biological questions and leads to the collection of samples that are needed to answer those questions. I have sectioned the experimental design elements, which deal with “Define” as the first step of the workflow, as follows:

(1) Type of experiment: Intervention study versus a real-world situation

For the purpose of this study, mothers who were HIV positive on an antenatal ART PMTCT programme were recruited to investigate whether the HIV and ART exposure of the embryo *in utero* had an effect on fetal homeostasis at birth. As mentioned above, this PhD was designed to use metabolomics as an investigative tool to review the homeostasis of neonates at birth.

The ambitious goal set by the Joint United Nations Programme on HIV/AIDS and the President’s Emergency Plan for AIDS Relief is to cut the number of new infections by 90% from the 2009 baseline. Consolidated guidelines such as the 2013 WHO guidelines recommend that all pregnant and breastfeeding women be administered triple therapy either from Option B or B+ (as defined in Chapter 2), and have largely impacted PMTCT programmes (WHO, 2013). The elimination of single-drug regimens (no longer recommended by WHO) and the global adoption of Option B+, leading to longer duration exposures of ARVs during pregnancy and breastfeeding, including ART for ineligible pregnant/breastfeeding women under Option B, has shifted the paradigm of PMTCT (Kellerman *et al.* 2013). The new era of “test and treat” offers women combination (or combined) ART (cART), which has become the global standard for PMTCT. In Option B the recommended prophylaxis regimen includes: AZT + 3TC + LPV/r; AZT + 3TC + ABC; AZT + 3TC + EFV; TDF + 3TC (or FTC) + EFV and the recommended Option B+ regimens are: TDF + 3TC (or FTC) + EFV (see Chapter 2, Table 2-2). The majority of HIV-infected women meet the criteria of antiretroviral treatment with two reverse transcriptase inhibitors and a protease

inhibitor (thereby representing cART). With regard to my study, published in *Metabolomics* in 2016 (see Chapter 4), the mothers were administered different treatment regimens as seen fit by their clinicians. This project was thus not designed as an intervention study but rather the outcome of sampling in a real-world scenario. John P. A. Ioannidis (2016) describes “*clinical research*” to be useful when “*it can lead to a favourable change in decision making (when changes in benefits, harms, costs and any other impact are considered) either by itself or when integrated with other studies and evidence in systemic reviews, meta-analyses, decision analysis and guidelines*”. He further articulates this in his paper “Why most clinical research is not useful”. Experimental intervention studies (clinical trials) are meant to address questions regarding treatment and prevention and are often compared to giving a placebo or normal care. Therefore, there is a group of participants randomly assigned either to a control or experimental group, where the experimental group would receive the treatment and the control group would not receive treatment (or placebo). Experimental intervention studies are often designed with the aim of either supporting or discrediting a hypothesis, therefore making this form of research deductive (Gray 2013). In real-world research, a true experimental research design is not possible since it would be virtually impossible to carry out.

Maternal-fetal investigations such as this one have to take into consideration not just the effect the mother’s condition and details will have on the study, such as age, race, diet and lifestyle – for example, smoking and drinking habits – but also the infant’s health criteria, such as Apgar scores, gestational age, birth weight and cord blood pH. In a real-world scenario, these factors are not under prospective control. In an ideal metabolomics study, research involving assessing the mechanisms of a diseased group versus a control group (that is, uninfected subjects), the recruited participants would be the same age, gender, race, diet and exposed to the same lifestyles circumstances, which is not the case in the real world, and was the scenario in which my investigation was rooted.

(2) Experimental case descriptions: The imperative for participation of clinicians

In our metabolomics investigation, comprehensive information on the individual, environmental and clinical profile of each participating experimental subject (mother and child) and consistency in sample collection, description and storage were essential and a prerequisite not only to achieving a successful metabolomics study, but also to enable the identification of potential confounding effects which can affect such an investigation (Dunn *et al.* 2012). As a biochemist, I do not possess the skills to collect

cord blood and to make clinical assessments, among other talents needed; therefore, the participation of a qualified clinician who adheres strictly to standard operating protocols when collecting cord blood was critically important in my study. At birth, immediately after the infant is delivered, the umbilical cord is clamped and cord blood is collected prior to the umbilical cord being cut. A clinician is also required to correctly label the collected cord blood at the time of collection along with details such as collection date, time and patient ID. Study material for this investigation was collected from the Emma Children's Hospital at the Academic Medical Centre (AMC) of the University of Amsterdam in the Netherlands. Likewise, all details of the mothers had to be recorded, as will be described in Chapter 4 (see published manuscript)

Johnson and co-workers (2016) recommend establishing experimental designs using "*patient questionnaires with subsequent population stratifications*", along with incorporating statistical methods such as regression modelling to narrow in on extractions of significant metabolites (Ellis *et al.* 2012) and removing confounding samples from the analysis. A clinician who is involved in patient care of both mother and child, would typically be handling patient questionnaires and making provisions for a longitudinal study if the study sees fit. In their review Dunn *et al.* (2012) defines an experimental design as "Design of a study to acquire data related to a specific biological question while ensuring that covariants or confounders are not present or are well characterized".

(3) Ethical approval

Before an experiment can even be contemplated or conducted, it is of the utmost importance that medical research investigations acquire the necessary ethical approval and that all contributors involved in the investigation offer their consent. An article by Bauchner and Sharfstein reviews published papers which have failed to report ethical approval in child health research (Bauchner & Sharfstein 2001). In order for researchers to publish their work in respectable journals, it is now required by over 500 journals that have adopted the guidelines of the International Committee of Medical Journal Editors that the authors should "*indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation*" (Rennie 1997). Research on child-related investigations involving patients younger than 18 years, pregnant woman and child care by clinicians, published in 1999 in five American journals, reported that, despite concerns regarding the protection of human subjects, 40% of these published articles did not report ethical clearance even with the journal's explicit requirement to do so (Bauchner & Sharfstein

2001). More surprising was that a quarter of these medical journals do not even require the confirmation of ethical clearance from their authors (Amdur & Biddle 1997). Ethical approval for studies involving human participants are important as it conveys a message to the reader that the human subjects involved were treated in accordance with contemporary ethical standards (Kanter 2009). Since the start of civilization, ethics in medicine has been practised from the time of the introduction of the Hippocratic oath formulated as – “*thou shall do no harm*” – which was introduced in the 4th century BC when Hippocrates himself was establishing medicine as a science, advancing it from its primitive state (Resnik 2011). The Declaration of Helsinki oath (1964), sets recommendations in research to guide physicians in biomedical research involving human subjects (World Medical Organization. Declaration of Helsinki, 1996). In a 2011 review, Resnik offers a comprehensive outline of what ethics is and why it should be used. He also outlines that different professional associations, government agencies, universities and research centres require good ethics in order for research to be conducted and so do many other governmental agencies such as the US National Institutes of Health (NIH), the National Science Foundation (NSF), and the Food and Drug Administration to name a few (Resnik 2011).

In my investigation, ethical approval was granted for both my controls and the use of cord blood samples collected from the Emma Children’s Hospital of the AMC (see Chapter 4, and Annexure A1) and the ethical code of the Potchefstroom Laboratory for Inborn Errors of Metabolism (PLIEM), hosted at North-West University (Potchefstroom Campus) for the use of infant control serum samples.

(4) Biofluid: Cord blood

In human-centred metabolomics studies the choice of sample is based on the objective of the study. A wide range of biological material, such as tissues, cells and biofluids usually composed of blood plasma, serum, urine or cerebrospinal fluid, is generally analysed. With regard to metabolomics application in maternal-fetal medicine biological samples can be derived from (1) the mother: urine, plasma, vaginal fluids, saliva and breast milk; (2) the fetus: amniotic fluid, umbilical cord blood; and (3) the infant: urine, plasma, placental tissue and saliva (Fanos *et al.* 2012).

To our knowledge, I am not aware of any studies involved in the metabolomic analysis of HIV-ART exposure in neonates and linking umbilical cord blood metabolites to the effects of such exposure *in utero*. We are, however, aware that the cord blood metabolome is exceedingly informative and has been associated with important clinical

outcomes, providing better understanding not only of fetal HIV-ART exposure but also pathogenesis of different diseases. In Chapter 2, I outline the effects of an infant exposed *in utero* to HIV-ART and raise concerns regarding the growing cohort of these infants. The investigative analysis of the cord blood metabolome of a neonate exposed to HIV-ART *in utero* is, however, fully discussed in Chapter 4.

The umbilical vein, supplies the fetus with oxygenated, nutrient-rich blood from the placenta and reflects the maternal, acid base status and placental function. Umbilical arteries are smaller in size and present a window on the status of the neonate, conversely and simultaneously transport deoxygenated nutrient-depleted arterial blood back to the placenta (Thorp *et al.* 1996). In metabolomics there is a great range of reviews and investigations, such as those by Emwas *et al.* (2013) and Bernini *et al.* (2011), that offer extensive knowledge on urine and plasma sample collection, storage and analysis for NMR sample analysis; however, there is little information on cord blood sample collection in metabolomics. Moreover, once collected in Eppendorf tubes or vacutainers and allowed to clot (in the case of serum) or be centrifuged (if plasma), the experimental procedure follows the same plasma/serum storage protocols (Bernini *et al.* 2011). Urine is famously used in metabolomics investigations because it encompasses a broader spectrum of the human metabolome in a biological system. It is obtained non-invasively and can be collected in large volumes. However, urine is unable to provide any information regarding the prenatal metabolite transfer or detailed metabolite profile of the maternal-fetal interaction at birth (Tea *et al.* 2012). Cord blood collected for metabolomics analysis can be stored for a short term at -20°C (Tea *et al.* 2012) until required and for a longer term at -80°C (Bernini *et al.* 2011). Since the objective of our study was to investigate the maternal-fetal effect of HIV-ART exposure and measure the effects at birth, cord blood is the only biological sample which satisfactorily addresses our biological question. It offers intrauterine metabolic information on both the maternal-fetal relationship and it is non-invasively obtained.

There is a growing body of investigations that employ the use of cord blood in their neonatal clinical investigations, such as the one presented here for my PhD. These investigations, whether predictive or diagnostic/prognostic biomarker discovery studies, use biobanks to store their collected umbilical cord samples (O'Donovan *et al.* 2014). Umbilical cord blood studies can use either arterial or venous cord blood. In a study by Paskova *et al.* in 2014, however, it was recognized that the use of cord blood in preterm babies is not practical as sample collection becomes a tedious and difficult exercise. There is a large difference in the diameters of the umbilical cord at 23 weeks

gestational age and that of 40 weeks gestational pregnancy. In preterm babies it is almost impossible to separate the arterial and venous cord blood (Paskova *et al.* 2014). Tea and colleagues, however, in their ¹H-NMR-based metabolic study which profiled the umbilical vein and arterial cord blood of very low birth weight preterm infants, were able to record differences based on altered maternal-fetal nutrient exchange compared to a control group (Tea *et al.* 2012).

There is a prevalence of 3.3% haemolysis in routine venepuncture sample collections (Lippi *et al.* 2008). However, this statistic is unknown for collections of bio-banked umbilical cord blood. Denihan and co-workers (2015), in their metabolomics investigation, assessed the effect that *in vitro* haemolysis could have on metabolomic alterations in umbilical cord blood samples, a common procedural problem which impacts umbilical cord blood sampling. Denihan *et al.* (2015) presented retrospective data from two published metabolomics-based experimental investigations: (1) Walsh *et al.* (2012): “*The metabolomic profile of umbilical cord blood in neonatal hypoxic ischemic encephalopathy*”, and (2) Reinke *et al.* (2013): “*¹H-NMR derived metabolic profile of neonatal asphyxia in umbilical cord serum: implications of hypoxic ischemic encephalopathy*”. The purpose of these studies was to determine if visual haemolysis of the cord blood serum would affect metabolite concentrations in the healthy control population. They comprised 69 samples of cord blood that were classified as haemolysed against 56 samples without signs of haemolysis. All 69 samples were analysed using a combined targeted direct injection and LC-MS/MS method; the 59 samples (10 had insufficient volume) were analysed by ¹H-NMR. Using these methods, the authors found 43 metabolites to be significantly altered when comparing the haemolysed with the clean serum samples. In these results there was evidence of changes in metabolite concentrations within samples with visual haemolysis; these included acylcarnitine (fold change, FC) increase of 1.2), and glycerophospholipids (FC decrease of 1.3). Metabolites associated with the Krebs cycle: acetate (FC increase 1.5), formate (FC increase 1.8), succinate (FC increase 1.9) and the non-essential amino acid ornithine (FC increase 1.5) were a few of the results presented. In their discussion of their retrospective study, the authors (Denihan *et al.* 2015) examined each metabolite to determine if the observed changes were due to metabolite degradation or if they were metabolized by enzymes released from the haemolysed erythrocytes (Yin *et al.* 2013). Increases from the significantly increased metabolites included acylcarnitines (L-acetylcarnitine and hexanoylcarnitine), a sugar glycerol, an amino acid (phenylalanine), a non-essential amino acid (ornithine), and Krebs cycle intermediates (citrate, succinate, acetate and formate). The change of

arginine to ornithine was associated with the erythrocytes and monocytes released into the cytoplasm. Another study, by Theil *et al.* (2012), indicated in an animal study on how haemolysis can increase acetate and decrease glucose in metabolites examined from plasma samples; this is also true for falsely recorded creatinine in clinical studies (Lippi *et al.* 2008). To conclude, Denihan and co-workers, in a first metabolomics study to compare umbilical cord blood metabolites of haemolysed serum samples from healthy term infants, agree that although the removal of these haemolysed samples for biomarker discovery is not practical, a library is preferable to identify metabolites that are unreliable to measure in haemolysed samples (Denihan *et al.* 2015).

The above background provides a strong indication of how meticulous an analyst and the appointed clinician need to be when handling and collecting metabolomics samples. If not handled correctly, there can be the consequences of reporting unreliable, even false data. In my own study, no haemolysis of samples was observed, and the cord blood was obtained from venous cord blood, with virtually no contamination from arterial blood. Further details on the importance of cord blood in clinical neonatal metabolomic studies are reviewed and discussed in Chapter 7.

3.4.2 Acquire: Sample processing and measurement

Metabolomic investigations involving human subjects are generally assigned to three categories, as reviewed by Dunn *et al.* (2012):

- (1) Pathogenesis studies, which are compiled with the goal of uncovering mechanisms of disease;
- (2) Predictive biomarker discovery, which aims to identify a metabolite signature that correlates a disease and environment with the object of being translated into diagnostic or prognostic tests;
- (3) Association studies, which are intended to discover relationships between the human metabolome and factors such as physical activity, demographics and lifestyle.

In a metabolomics experiment it has been established that there is no single extraction method or biofluid that is ideal for measuring all the metabolites within an entire organism (as extensively reviewed by Gebregiworgis and Powers for NMR-based studies). With this said, it should be noted that the extraction method of sample choice and time of sampling can cause large variations and influence the reproducibility of the sample analysis (Gebregiworgis & Powers 2012). The strength of a metabolomics

system lies not in its ability to detect multiple signals but rather in its strength to identify multiple metabolites. And so in metabolomics data, its “*quality over quantity*”, since a single signal that is identifiable is more informative and valuable than a mass of unidentified low-level signals commonly known as noise (Verpoorte *et al.* 2008).

Most analytical platforms for metabolomic studies use targeted, semi-targeted and untargeted strategies (Dunn *et al.* 2011).

- **Untargeted metabolomic investigations (NMR)**

Details of my NMR data are presented in Chapter 4. Spectral regions devoid of peaks (noise) and the region containing water resonances were excluded from my data. All data were normalized in the AMIX software program by dividing each integrated segment by the total spectrum intensity (minus the excluded regions), which compensated for the differences in overall concentration between individual plasma samples. Assignment of potential metabolites to each bin resulted in 54 bins with each corresponding to a distinct metabolite. The remaining bins contained substances and features which could not be annotated unequivocally as distinct metabolites. Statistical analysis revealed that the 54 bins contained meaningful information on transient neonatal hypoglycaemia and on fetal HIV-ART exposure, which was not observed for the excluded bins (Annexure A1-2, A1-3).

- **Semi-targeted metabolomic investigations (GC-MS and LC-MS)**

These studies are designed to provide semi-quantitative concentrations of predefined metabolites. This approach, unlike the targeted analytical method, involves a holistic study of the metabolome and is capable of detecting intermediary metabolites and biochemicals (Dunn *et al.* 2011; Zhou *et al.* 2012). I include examples of semi-targeted studies of cord blood samples from my own experimental group, as well as from a comparative but not identical group in Chapter 6.

- **Targeted metabolomic investigations (NMR and MS)**

Although the exact number of metabolites in a human metabolome is unknown, targeted analysis, which measures a limited number of specified metabolites, which range up to dozens in number, is the most developed analytical approach in metabolomics (Shulaev *et al.* 2006, Carraro *et al.* 2009). Targeted analytical studies mostly apply mass spectral (MS) platforms with high levels of specificity, making sure the analytical method developed properly measures its concentration

of absolute amounts for each metabolite (Dunn *et al.* 2012). This approach was not employed in the present study.

When choosing an analytical technique, one needs to bear in mind that a fundamental issue in metabolomics research, apart from consistency and reproducibility in sample handling, is the acquisition of the high-throughput analytical data, and uni- and multivariate data analysis. The most commonly used platforms for detection and measurement of metabolites in metabolomics is untargeted metabolomic studies using nuclear magnetic resonance spectroscopy (NMR) or the hyphenated mass spectroscopy techniques such as gas chromatography (GC) and liquid chromatography (LC) (Dunn *et al.* 2012).

With regard to the framework of our investigation for this PhD, inconsistency is most often seen in the application of MS analysis, which implies that the inclusion of NMR analysis is an added value for the assessment of possible metabolic pathways affected by HIV-ART exposure of the neonate. Liquid chromatography provides an in-depth assessment of complex metabolite mixtures with highly sensitivity and high-level structural (MS) information on metabolites. Together these two techniques complement each other and provide critical information on the metabolism affected by HIV-ART exposure.

- **Nuclear magnetic resonance (NMR) spectroscopy**

NMR is a high-throughput fingerprinting technique which produces rapid, comprehensible and robust metabolic screening. It is a non-destructive, non-biased method which requires little to no sample preparation and is capable of screening a broad range of metabolites (Wishart 2008). NMR has been widely used to acquire the metabolome in a diverse group of studies; in medicine it has been used to analyse challenges presented to the human such as diet, differences of the homeostasis in illnesses and growth among two or more classes of individuals. Its greatest disadvantage, however, is its sensitivity. NMR is famously recognized for its inability to analyse a volume of metabolites which are present at low concentrations. Chapter 4 gives a full review on the advantages of using NMR as an analytical tool.

- **Liquid chromatography–mass spectrometry (LC-MS)**

Mass spectrometry is celebrated for its high sensitivity and extensive use in metabolomic studies (Glinski & Weckwerth 2006). Mass spectrometry is used widely for metabolic fingerprinting and metabolite profiling. When coupled to a high-resolution

gas chromatograph or to the increasingly used liquid chromatography technique, it presents a rapid ability to analyse large numbers of metabolites. A disadvantage of GC-MS is that it often requires derivatization of the collected sample, which makes it biased against non-volatile, high-molecular-weight metabolites (Dunn & Ellis 2005). It is able, however, to simultaneously profile several hundred chemically diverse compounds such as organic acids, amino acids, sugars, sugar alcohols, aromatic amines and fatty acids (Jonsson *et al.* 2005). LC-MS, another system combined with mass spectrometry, differs from GC-MS in various ways. It also has high sensitivity and requires no derivatization, which simplifies sample preparation and its range in accommodating analyte polarity and molecular mass is wider than that of the GC-MS technique (Dunn & Ellis 2005). Although able to analyse a large sample, LC-MS has the shortcoming of being slow, which is a handicap when performing long analytical runs.

3.4.3 Discover: Data pre-treatment metabolite annotation

Once metabolomics data have been generated, the next phase focuses on gaining information from the data – generally described as the discovery phase. Metabolomics plays a significant role in the discovery of biological indicators, such as risk factors, biomarkers and biosignatures. Some metabolites may not be directly associated with a specific disease state, but may be indicators of a particular disease; for example, cholesterol and triglycerides are indicators of coronary heart disease (Austin 1989). Such metabolites are called metabolic risk factors or indicators. A metabolic biomarker may be described as a metabolite that is objectively measurable, evaluated and validated as an indicator of normal biological processes, pathogenic processes, or of a response to a therapeutic intervention, such as glucose for diabetes (Caveney & Cohen 2011). The third entity, a biosignature, comprises a number of related metabolites that could be indicators of a disease state – for example, methylcitric, 2-ketoglutaric, quinolinic and 4-hydroxyhippuric acids have been proposed as a biosignature for tuberculosis meningitis in a recent study conducted in our laboratory (Mason *et al.* 2016).

The specific research question in my investigation was to determine whether any HIV-ART-induced metabolic perturbations are discernible at birth in neonates exposed to HIV-ART (Chapter 2) The approach to address this question is basically exploratory by following the scientific method of induction (Kell & Oliver 2004), as discussed in 3.4.4. below. The aim of the investigation is, however, to address the concern following treatment for PMTCT by exploring metabolic profiles in the exposed neonates, which

is in essence a path of discovery. In metabolomics discovery studies, the key issue is how to extract relevant information from the metabolomics data – that is, how to identify relevant (in terms of biology) and significant (as deduced by statistics) metabolic profiles from the biofluid or other biological material derived from the experimental subjects. Irrespective of the method of data generation, all data produced in metabolomics studies are highly multivariate, which requires statistical methods to uncover the trends or significant information from the data, i.e. to identify meaningful biological information (quantified metabolites). Once this information is obtained, the next phase in the discovery path is interpretation, as is discussed below.

Two main methodological approaches are mostly followed to handle the complex metabolomics data sets: (1) data processing and pre-treatment and (2) data analysis. The ultimate aim of these statistical processes is knowledge discovery, which is well covered in a review by Boccard and co-workers (Boccard *et al.* 2010). Although this review focuses on MS data, the generic aspects covered in the paper apply equally to NMR-based data.

Boccard and co-workers indicate that data processing and pre-treatment entail statistical methods such as filtering (removal of spectral noise), feature detection [the binning of NMR spectral information or alignment of the mass-to-charge ratio (m/z) information of MS], normalisation (the removal of bias in the data, while preserving all biological information), and finally scaling (the application of statistical models to estimate numerical factors among samples, e.g. median intensities for adjustments to fit the models used for data analysis – the identification of important metabolites). All these methods require expert statistical skills which are beyond the requirements for a biochemical study, but which were provided for the study by statistics experts within the Centre for Human Metabolomics at North-West University. Most of these techniques were, however, used in all experimental sections presented in this thesis, and will be covered in the methodological parts of the relevant chapters.

Data generation yields spectral (NMR or MS) information that needs to be translated into biological information. The nomenclature in this regard is: (1) *Features*. The basic aim of data processing is to transform raw data files in such a way that facilitates easy access to characteristics of each observed MS ion or NMR spectral point. This basic information is called *features*, and are m/z values and retention times of the ions and ion intensity measurements in MS data, and NMR spectral ppm values and peak intensities. According to these definitions, features do not yet convey any biological information. (2) *Annotation*. At some appropriate stage in the process of data

generation or processing, the analytical information of the features should be translated into biological information. A dual approach underpins this strategy: (i) The strategy mostly starts with tentatively annotating metabolomic MS mass spectra or NMR spectral ppm values to compounds that match the elemental composition of small molecules found in publicly available databases. (ii) Ideally, the list of tentative annotations could then be further confirmed (by a process of identification) by additional physicochemical information such as determination of MS/MS fragmentation patterns or proton NMR chemical shift characterization (Kind & Fiehn 2006). Most metabolomics studies suffice with the first step (annotation), given the high quality of the present data basis for metabolite identification. (3) *Variables*. This is a generic term used in metabolomics studies to indicate metabolites (annotated or structurally verified) or features that could not be identified as metabolites, and are normally indicated as unknowns.

The outcome of data processing, pre-treatment and annotation is presented in the form of a metabolomics data matrix $X_{(n \times p)}$:

$$\begin{matrix} X \\ (n \times p) \end{matrix} = \begin{bmatrix} X_{1-1} & X_{1-2} & \cdots & X_{1-p} \\ X_{2-1} & X_{2-2} & \cdots & X_{2-p} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n-1} & X_{n-2} & \cdots & X_{n-p} \end{bmatrix}$$

The data matrix consists of rows of metabolites (or spectral information) designated by X_{1-1} to X_{1-p} (for case 1) and columns of the experimental cases, designated as X_{1-1} to X_{n-1} (for metabolite or spectrum 1). The matrix will be discussed in more detail in Chapter 5; it provides the basis for data analysis, the next key component of the discovery phase. Although often a single or very few highly reliable metabolites can be sufficient to serve as a risk indicator or biomarker, an extended set of variables/metabolites is mostly desired for the construction of a metabolic network or (global) metabolite profile for the biological issue under examination. Changes in metabolite levels may be large (as in inherited metabolic diseases) or subtle (as in affective disorders such as depression). Identification of such changes from a multivariate metabolomics matrix can only be achieved by statistical processing to determine the relevance of an observed change. In the present investigation, the metabolomics data were analysed by the two most commonly used multivariate modes

of analysis: (1) unsupervised principal component analysis (PCA) and (2) partial least squares discriminant analysis (PLS-DA).

PCA [originally developed more than a century ago by K. Pearson (Philos. Mag. 1901, 2, 559–572)] is generally and most commonly used as the starting point of multivariate metabolomics data analyses. PCA is a form of transformation of multivariate data within the matrix, and is mostly used for exploratory analyses by displaying systematic variations within a study group of cases. The first step of data analysis is to explore and discover the overall structure of the data, find trends and groupings in the data, e.g. differentiation (or not) between controls and affected cases. This stage of data analysis is not based on the identification of groupings in the experimental subjects, and is called a blind or unsupervised method. Such a method allows for an unbiased view of the data, i.e. it provides an outcome that does not assume any prior knowledge. Apart from PCA, another unsupervised method includes hierarchical cluster analysis – both methods are used in the present study.

PLS-DA is a supervised method of analysis, as the method is informed by the classification of each case in terms of a specific group, e.g. as a control or a patient (Smolinska *et al.* 2012). Supervised techniques make use of *a priori* known structures. This technique is particularly suitable for situations where fewer observations (i.e. number of experimental subjects) than measured variables (e.g. metabolites) are present, as is mostly the case in a metabolomics matrix. A limitation of supervised methods is that they construct models which fit the data perfectly, that is, by over-fitting (Westerhuis *et al.* 2008), even if there is no real relation between groups within the data set. Therefore, validation of the model is an important step in the interpretation of PLS-DA information. The most important advantage of PLS-DA is that it provides information about those variables that indicate differences between two or more classes, which are called variables important in projection (VIP). PLS-DA is therefore a powerful approach in metabolomics biomarker discovery studies, and is applied extensively in the present study.

3.4.4 Interpretation: From information to hypothesis formulation to betterment

The chief objective of a metabolomics investigation is to translate raw data obtained from the selected analytical instrument into meaningful biological knowledge. Original views on metabolomics research tended to call it “*fishing experiments*” (Brent & Lok 2005). This perception was well refuted by several methodological/perspective papers

by Douglas Kell and his co-workers (Kell 2007; Kell & Oliver 2004). A key aspect of knowledge production, emphasized by Kell, was the cycle between induction (hypothesis generation) and deduction (hypothesis testing), as shown in Figure 3-4.

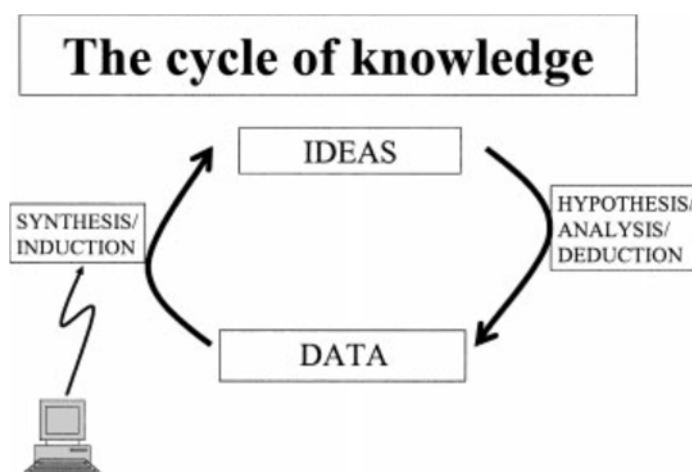


Figure 3–4: Scientific advance may be seen as an interactive cycle linking knowledge and observations. (Reproduced with permission from Kell & Oliver 2004).

An important point of departure in Kell and Oliver's paper (Kell & Oliver 2004) is that inductive modes of research (metabolomics: from data to ideas) do not compete with the deductive mode (hypothesis testing: from idea to data), but are complementary scientific approaches. Knowledge about metabolism has grown for more than a century through deductive research: The enzymes, cofactors and coenzymes, substrates, intermediates and products of individual biochemical reactions and the multitude of molecular pathways comprising intermediary metabolism are more completely understood than probably any other aspect of human biology. This comprehension of knowledge has been a major asset in understanding human health and improving the human condition as a resource for the development of foods, drugs, agricultural chemicals and in targeting severe disease conditions, by improving diagnosis, treatment and prognosis. Interest in metabolism as a field of research waned, however, in the glamour of molecular biology, following the elucidation of the structure of DNA in the early 1950s.

With the advent of metabolomics, however – linked to the other omics sciences such as genomics and proteomics – things have changed. Integrated metabolite data and their link to a distinct metabolic pathway do not suffice for a comprehensive understanding of metabolism. To achieve this, metabolomics has added a new

dimension by providing novel holistic insights in metabolism and on individual metabolic variability and identity, derived from a holistic understanding derived from a suit of methods. More precisely:

- The holistic approach takes an opposite view to the reductionist, hypothesis testing methodology as it is directed towards a system as a whole in which the complexities and interactions in the intact system come into focus.
- One fundamental difference of metabolomics from traditional metabolite research is that in using tailor-made analytical technologies, it measures quantitatively all (or as many as possible) analytes, in contrast to traditional (clinical) chemistry in which one or only a few metabolites are measured.
- A now-common strategy in post-genomic biology, including metabolomics, is the stimulation of the development of an innovative arsenal of analytical chemistry instruments, data-mining devices, mathematical algorithms and computer software to assist in the generation, identification and quantification of compounds towards the lower limit of detection ranges.
- Single biomarker-based diagnostics for metabolism-based diseases fail to identify the underlying reason why the surrogate (diagnostic biomarker) became altered – a limitation that metabolomics strives to overcome using accurate, actionable information about how to intervene in the development of metabolic disease process.
- A course of action to modulate perturbed metabolism within a distinct pathological trajectory may show variability within the individual context, which requires identification of the basis of the dysregulation. Thus, more than single markers of disease must be measured to provide the individual metabolic context, to enable individual specific diagnoses and prescriptions and to open the progressive need for personalized medicine as a counterpart to population-based medicine. Metabolomics is probably better suited for determining individual metabolic context, which in the end is the outcome of an integration of individual genetic, nutritional, pharmacological and environmental status.

Against this background it is clear that the interpretation of metabolomics information entails multidisciplinary cooperation and is much more than only hypothesis generation – important though that may be. Exploratory metabolomics is *par excellence* a route of discovery, aimed at the eventual betterment of individuals and society as a whole. Thus, the outcomes of metabolomics research are ultimately not intended to be put

aside and ignored as new knowledge, but are aimed to be translated and applied to achieve its goal of human betterment in health and disease.

3.4.5 Disseminate: New/translational knowledge

What sets metabolomics apart from the other omics sciences and is commonly perceived as its strength is that when conducting a metabolomics investigation, such as metabolite fingerprinting, one does not need to have prior knowledge of the complex metabolites of a pathological condition that might be discovered, leading to potential biomarker discovery that were not even considered (Fanos *et al.* 2015). However, it should not be overlooked how far other omics platforms have come. In the *New England Journal of Medicine* Reza Mirnezami and co-authors discuss the potential of “*precision medicine*” – this widely researched concept deals with “*establishing clinical-pathological indexes with state of the art molecular profiling to create diagnostic, prognostic and therapeutic strategies precisely tailored to each patient’s requirements*” (Mirnezami *et al.* 2012). In order for precision medicine to be more than just a concept, the regulatory bodies that regulate clinical trials need to evolve at the same pace as biotechnical advances. In neonatal medicine where there is an urgent need for biomarkers for critically ill newborns with either neonatal sepsis, which commonly originates from an abrupt evolution of infections in the first four weeks of life (Mussap *et al.* 2015), or acute kidney injury, delays regarding study approval can lead to increases in neonatal mortality. According to Buyse *et al.* (2010), “Biomarkers used in preclinical drug development are identified as pharmacokinetic or pharmacodynamics biomarkers. When they predict the likely course of the disease irrespective of treatment, they are called prognostic biomarkers and finally when they forecast the response to treatment they are predictive”. In a review by van der Greef *et al.* (2006), “*Metabolomics-based system biology and personalized medicine: moving toward n = 1 clinical trials*” the authors believe that new advanced technological developments in metabolomics will be significant in the near future, analysis of metabolites will be analysed by fully automated systems which will be able to document thousands of metabolites with minimal sample volumes (5–10 µL) required (van der Greef *et al.* 2006). This will revolutionize phenotyping. In a review in *Biomarkers in Neonatology*, Mussap and Fanos (2015) state that “*metabolomics has become revolutionary if compared to traditional techniques. The evolution that has become revolution*”. Mussap and Fanos further believe that a future is approaching in which the clinical application of systems biology, proteomics and metabolomics will be more practical

and that applied science will play a larger role in improving health care and the survival of critically ill babies.

Metabolomics in my research has placed my experimental results in a more informative context than hitherto, by exposing the metabolite changes which reflect the surrounding environment of the fetus, *in utero* before birth and then immediately after birth. The outcome of an untargeted approach in this regard is presented in the next chapter.

REFERENCES

- Amdur, R. J., & Biddle, C. (1997). Institutional review board approval and publication of human research results. *JAMA*, 277(11), 909–914.
- Atzori, L., Antonucci, R., Barberini, L., Griffin, J. L., & Fanos, V. (2009). Metabolomics: a new tool for the neonatologist. *The Journal of Maternal-Fetal & Neonatal Medicine*, 22(Supplement 3), 50–53.
- Austin, M. A. (1989). Plasma triglyceride as a risk factor for coronary heart disease. The epidemiologic evidence and beyond. *American Journal of Epidemiology*, 129(2), 249–259.
- Barker, D. J. (1990). The fetal and infant origins of adult disease. *British Medical Journal*, 301, 1111.
- Barker, D. J., Eriksson, J. G., Forsén, T., & Osmond, C. (2002). Fetal origins of adult disease: strength of effects and biological basis. *International Journal of Epidemiology*, 31(6), 1235–1239.
- Bernini, P., Bertini, I., Luchinat, C., Nincheri, P., Staderini, S., & Turano, P. (2011). Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. *Journal of Biomolecular NMR*, 49, 231–243.
- Bino, R. J., Hall, R. D., Fiehn, O., Kopka, J., Saito, K., Draper, J., et al. (2004). Potential of metabolomics as a functional genomics tool. *Trends in Plant Science*, 9(9), 418–425.
- Boccard, J., Veuthey, J. L., & Rudaz, S. (2010). Knowledge discovery in metabolomics: an overview of MS data handling. *Journal of Separation Science*, 33(3), 290–304.
- Brent, R., & Lok, L. (2005). A fishing buddy for hypothesis generators. *Science*, 308, 504–506.
- Buyse, M., Sargent, D. J., Grothey, A., Matheson, A., & De Gramont, A. (2010). Biomarkers and surrogate end points—the challenge of statistical validation. *Nature Reviews Clinical Oncology*, 7(6), 309–317.

- Carraro, S., Giordano, G., Reniero, F., Perilongo, G., & Baraldi, E. (2009). Metabolomics: a new frontier for research in pediatrics. *The Journal of Pediatrics*, 154, 638–644.
- Caveney, E. J., & Cohen, O. J. (2011). Diabetes and biomarkers. *Journal of Diabetes Science and Technology*, 5(1), 192–197.
- Collins, F. S. (2011). Re-engineering translational science: the time is right – Commentary. *Science Translational Medicine*, 3, 1–9.
- Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine*, 372(9), 793–795.
- Denihan, N. M., Walsh, B. H., Reinke, S. N., Sykes, B. D., Mandal, R., Wishart, D. S., et al. (2015). The effect of haemolysis on the metabolomic profile of umbilical cord blood. *Clinical Biochemistry*, 48(7), 534–537.
- Dessi, A., Atzori, L., Noto, A., Adriaan Visser, G. H., Gazzolo, D., Zanardo, V., et al. (2011). Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolic syndrome. *The Journal of Maternal-Fetal & Neonatal Medicine*, 24(Supplement 2), 35–39.
- Dunn, W. B., & Ellis, D. I. (2005). Metabolomics: current analytical platforms and methodologies. *TrAC Trends in Analytical Chemistry*, 24(4), 285–294.
- Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R., & Griffin, J. L. (2011). Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews*, 40(1), 387–426.
- Dunn, W. B., Wilson, I. D., Nicholls, A. W., & Broadhurst, D. (2012). The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis*, 4, 2249–2264.
- Ellis, J. K., Athersuch, T. J., Thomas, L. D., Teichert, F., Pérez-Trujillo, M., Svendsen, C., et al. (2012). Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. *BMC Medicine*, 10, 1–10.

- Emwas, A. H. M., Salek, R. M., Griffin, J. L., & Merzaban, J. (2013). NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations. *Metabolomics*, 9(5), 1048–1072.
- Fanos, V., Atzori, L., Makarenko, K., Melis, G. B., & Ferrazzi, E. (2013). Metabolomics application in maternal-fetal medicine. *BioMed Research International*, 2013, 1–9.
- Fanos, V., Barberini, L., Antonucci, R., & Atzori, L. (2011). Metabolomics in neonatology and pediatrics. *Clinical Biochemistry*, 44(7), 452–454.
- Fiehn, O. (2002). Metabolomics—the link between genotypes and phenotypes. *Plant Molecular Biology*, 48, 155–171.
- Fiehn, O., Kristal, B., Ommen, B. V., Sumner, L. W., Sansone, S. A., Taylor, C., et al. (2006). Establishing reporting standards for metabolomic and metabonomic studies: a call for participation. *Omics: A Journal of Integrative Biology*, 10(2), 158–163.
- Gates, S. C., & Sweeley, C. C. (1978). Quantitative metabolic profiling based on gas chromatography. *Clinical Chemistry*, 24, 1663–1673.
- Gebregiworgis, T., & Powers, R. (2012). Application of NMR metabolomics to search for human disease biomarkers. *Combinatorial Chemistry & High Throughput Screening*, 15(8), 595–610.
- Glinski, M., & Weckwerth, W. (2006). The role of mass spectrometry in plant systems biology. *Mass Spectrometry Reviews*, 25(2), 173–214.
- Gluckman, P. D., & Hanson, M. A. (2004). Developmental origins of disease paradigm: a mechanistic and evolutionary perspective. *Pediatric Research*, 56(3), 311–317.
- Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *New England Journal of Medicine*, 359(1), 61–73.
- Gray, D. E. (2013). *Doing research in the real world*. Sage Publications, London.
- Hashimoto, F., Nishiumi, S., Miyake, O., Takeichi, H., Chitose, M., Ohtsubo, H., et al. (2013). Metabolomics analysis of umbilical cord blood clarifies changes in saccharides associated with delivery method. *Early Human Development*, 89(5), 315–320.

Holmes, E., Wilson, I. D., & Nicholson, J. K. (2008). Metabolic phenotyping in health and disease. *Cell*, 134(5), 714–717.

Ioannidis, J. P. (2016). Why most clinical research is not useful. *PLoS Medicine*, 13(6), 1–10.

Johnson, C. H., Ivanisevic, J., & Siuzdak, G. (2016). Metabolomics: beyond biomarkers and towards mechanisms. *Nature Reviews Molecular Cell Biology*, 1–9.

Jonsson, P., Bruce, S. J., Moritz, T., Trygg, J., Sjöström, M., Plumb, R., et al. (2005). Extraction, interpretation and validation of information for comparing samples in metabolic LC/MS data sets. *Analyst*, 130(5), 701–707.

Kanter, S. L. (2009). Ethical approval for studies involving human participants: academic medicine's new policy. *Academic Medicine*, 84(2), 149–150.

Kell, D. B. (2007). Metabolomic biomarkers: search, discovery and validation. *Expert Review of Molecular Diagnostics*, 7(4), 329–333.

Kell, D. B., & Oliver, S. G. (2004). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays*, 26(1), 99–105.

Kell, D. B., & Oliver, S. G. (2016). The metabolome 18 years on: a concept comes of age. *Metabolomics*, 12, 1–8.

Kellerman, S. E., Ahmed, S., Feeley-Summerl, T., Jay, J., Kim, M. H., Koumans, E., et al. (2013). Beyond PMTCT: Keeping HIV Exposed and Positive children healthy and alive. *AIDS*, 27(2), S225–S233.

Kind, T., & Fiehn, O. (2006). Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics*, 7(1), 1–10.

Lippi, G., Blanckaert, N., Bonini, P., Green, S., Kitchen, S., Palicka, V., et al. (2008). Haemolysis: an overview of the leading cause of unsuitable specimens in clinical laboratories. *Clinical Chemistry and Laboratory Medicine*, 46(6), 764–772.

Loscalzo, J., Kohane, I., & Barabasi, A. L. (2007). Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Molecular Systems Biology*, 3, 1–11.

- Mason, S., van Furth, A. M. T., Solomons, R., Wevers, R. A., van Reenen, M., & Reinecke, C. J. (2016). A putative urinary biosignature for diagnosis and follow-up of tuberculous meningitis in children: outcome of a metabolomics study disclosing host–pathogen responses. *Metabolomics*, 12(7), 1–16.
- Mirnezami, R., Nicholson, J., & Darzi, A. (2012). Preparing for precision medicine. *New England Journal of Medicine*, 366(6), 489–491.
- Moco, S., Collino, S., Rezzi, S., & Martin, F. P. J. (2013). Metabolomics perspectives in pediatric research. *Pediatric Research*, 73, 570–576.
- Mussap, M., & Fanos, V. (2015). Biomarkers in neonatology. *General Methods in Biomarker Research and their Applications*, 457–481.
- Mussap, M., Antonucci, R., Noto, A., & Fanos, V. (2013). The role of metabolomics in neonatal and pediatric laboratory medicine. *Clinica Chimica Acta*, 426, 127–138.
- Nicholson, J. K. (2006). Global systems biology, personalized medicine and molecular epidemiology. *Molecular Systems Biology*, 2(52), 1–8.
- Nicholson, J. K., Lindon, J. C., & Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, 29(11), 1181–1189.
- O'Donovan, S. M., Murray, D. M., Hourihane, J. O. B., Kenny, L. C., Irvine, A. D. & Kiely, M., (2014). Cohort profile: the Cork BASELINE Birth Cohort Study: babies after SCOPE: evaluating the longitudinal impact on neurological and nutritional endpoints. *International Journal of Epidemiology*, 44(3), 764–775.
- Oliver, S. G., Winson, M. K., Kell, D. B., & Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology*, 16, 373–378.
- Paskova, A., Parizek, A., Hill, M., Velíková, M., Kubátová, J., Duskova, M., et al. (2014). Steroid metabolome in the umbilical cord: is it necessary to differentiate between arterial and venous blood?. *Physiological Research*, 63(1), 115–126.
- Powers, R. (2015). Preface: “The whole is greater than the sum of its parts.”-Aristotle. *Current Metabolomics*, 3(1), 2–3.

- Reinke, S. N., Walsh, B. H., Boylan, G. B., Sykes, B. D., Kenny, L. C., Murray, D. M., & Broadhurst, D. I. (2013). ¹H NMR derived metabolomic profile of neonatal asphyxia in umbilical cord serum: implications for hypoxic ischemic encephalopathy. *Journal of Proteome Research*, 12(9), 4230–4239.
- Rennie, D. (1997). Disclosure to the reader of institutional review board approval and informed consent. *JAMA*, 277(11), 922–923.
- Resnik, D. B. (2011). What is ethics in research & why is it important. In *The National*, 1–10.
- Robertson, D. G., Watkins, P. B., & Reily, M. D. (2010). Metabolomics in toxicology: preclinical and clinical applications. *Toxicological Sciences*, 120(suppl_1), S146–S170.
- Shulaev, V. (2006). Metabolomics technology and bioinformatics. *Briefings in Bioinformatics*, 7(2), 128–139.
- Smolinska, A., Blanchet, L., Buydens, L. M., & Wijmenga, S. S. (2012). NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review. *Analytical Chimica Acta*, 750, 82–97.
- Tea, I., Le Gall, G., Küster, A., Guignard, N., Alexandre-Gouabau, M. C., et al. (2012). ¹H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*, 7(1), 1–12.
- Theil, P. K., Pedersen, L. J., Jensen, M. B., Yde, C. C., & Knudsen, K. E. (2012). Blood sampling and hemolysis affect concentration of plasma metabolites. *Journal of Animal Science*, 90(Supplement 4), 412–414.
- Van Der Greef, J., Hankemeier, T., & McBurney, R. N. (2006). Metabolomics-based systems biology and personalized medicine: moving towards n= 1 clinical trials?. *Pharmacogenomics*, 7(7), 1087–1094.
- Verpoorte, R., Choi, Y. H., Mustafa, N. R., & Kim, H. K. (2008). Metabolomics: back to basics. *Phytochemistry Reviews*, 7(3), 525–537.
- Walsh, B. H., Broadhurst, D. I., Mandal, R., Wishart, D. S., Boylan, G. B., Kenny, L. C., & Murray, D. M. (2012). The metabolomic profile of umbilical cord blood in neonatal hypoxic ischaemic encephalopathy. *PLoS One*, 7(12), 1–12.

Westerhuis, J. A., Hoefsloot, H. C., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J., et al. (2008). Assessment of PLS-DA cross validation. *Metabolomics*, 4(1), 81–89.

WHO. (2013). Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection World Medical Organization. Declaration of Helsinki. (1996). *British Medical Journal*, 313(7070), 1448–1449.

Wishart, D. S., Lewis, M. J., Morrissey, J. A., Flegel, M. D., Jeroncic, K., Xiong, Y., et al. (2008). The human cerebrospinal fluid metabolome. *Journal of Chromatography B*, 871(2), 164–173

Yin, P., Peter, A., Franken, H., Zhao, X., Neukamm, S. S., Rosenbaum, L., et al. (2013). Preanalytical aspects and sample quality assessment in metabolomics studies of human blood. *Clinical Chemistry*, 59(5), 833–845.

Zhou, B., Xiao, J. F., Tuli, L., & Ransom, H. W. (2012). LC-MS-based metabolomics. *Molecular BioSystems*, 8(2), 470–481.

CHAPTER 4 METABOLIC RISKS AT BIRTH FOLLOWING HIV-ART EXPOSURE: INSIGHTS FROM AN UNTARGETED METABOLOMICS STUDY

4.1 Introduction

4.1.1 Birth asphyxia and extra-uterine adaptation

The placenta is the key organ for fetal growth, immunological and endocrine functions and development. Moreover, it provides the interface between mother and fetus for the regulation of the fetal–maternal exchange of nutrients, oxygen and carbon dioxide, ions, and metabolic degradation products. In its intrauterine environment, the neonate thus develops to be fully dependent on the uninterrupted transplacental nutrient supply it obtains from its mother (Platt & Deshpande 2005). This dependence changes dramatically at birth. Because of normal uterine contraction during the birth process, all fetuses experience some degree of asphyxia – a condition of impaired gas exchange that leads to biochemical aberrations such as hypoxaemia and metabolic acidosis. Clinical assessment of the severity of asphyxia is difficult because the duration and nature of the exposure and the characteristics of the fetal cardiovascular response to the asphyxia in the affected fetus are usually not known (Low 1997). An infant in distress at full term requires immediate attention of the paediatrician or other medical personnel assisting the birth process as a fetus that experiences a significant asphyxia is at risk of developing hypoxic-ischaemic encephalopathy or other end organ sequelae.

Leuthner (2004), however, averred that a diagnosis of severe asphyxia should be made only if the neonate meets all conditions defined in the 5th edition of *Guidelines for Perinatal Care*: “(1) umbilical cord arterial pH less than 7 (whether metabolic or mixed), (2) Apgar score of 0 to 3 for longer than 5 minutes, (3) neurologic manifestations (eg, seizures, coma, or hypotonia), and (4) multisystemic organ dysfunction” (American Academy of Pediatrics, & American College of Obstetricians and Gynecologists 2002). They emphasize that it is important to provide proper resuscitation, to support infants, and to allow time for evaluation. According to the information provided on HIV-ART and the control infants used in this study, no indications were observed of serious or life-threatening asphyxia in any of the experimental cases (see in manuscript section 4.4.2).

4.1.2 Metabolic perturbations at birth and extra-uterine adaptation

Apart from a degree of asphyxia, disruption of the maternal–fetal interphase causes also metabolic challenges at and following birth, having distinct risks for the neonate. These metabolic changes include discontinuation of the prenatal assured glucose supply, which requires the mobilization of glucose from glycogen and fatty acids and triglycerides from lipids for adequate energy supplies (Hawdon *et al.* 1992, Cornblath *et al.* 2000). In order for the neonate to survive the stress experienced during the birth process and to adapt to its new extra-uterine environment, a well-regulated system such as the endocrine stress response, controlled under the influence of hormonal surges and timely expression of genes, drives metabolomic changes such as hepatic glycogenolysis, lipolysis, fatty acid β -oxidation, ketogenesis along with the creation of lactate and ketoacids derived from proteolysis (Cornblath *et al.* 2000). Although it has been indicated that at birth the plasma glucose concentrations recorded in the umbilical venous blood is initially 80–90% of the maternal venous blood (Heck & Erenberg 1987), there is a drastic decrease of glucose in the neonatal plasma in the first hour of birth, which is usually marked by hypoglycaemia (Hawdon *et al.* 1992). Subsequent regulation of endogenous production of glucose by the liver steadily increases glucose at a rate of 4–5 mg/kg/min (Platt & Deshpande 2005).

Neonatal hypoglycaemia is a common metabolic problem affecting as many as 3–29% of pregnancies (Johnson 2003), yet its definition of clinical significance remains controversial in contemporary neonatology (Cornblath 1997). A study by Harris, Weston and Harding noticed that, although routine blood glucose screening is recommended for neonates at risk for hypoglycaemia, the incidence of those screened is not well described and increasingly screening methods used to detect hypoglycaemia in neonates are inconsistent and unreliable (Harris *et al.* 2012). For glucose monitoring of clinical neonatal hypoglycaemia in neonates, thresholds have been used in clinical practice by clinicians for treatment and are recognized usually at <2.2 mM or <2.6 mM, for an estimated 5–15% of otherwise healthy babies (Cornblath *et al.* 2000). In their review “*Controversies regarding definition of neonatal hypoglycemia: Suggested operational thresholds*”, Cornblath, M., Hawdon, M. J., Williams, A. F., *et al.* (2000) acknowledge that there is no single measure that can define significant hypoglycaemia to be applied universally to every individual patient. They suggest monitoring of glucose concentrations as soon as possible after birth and

again some hours prior to and after feeding. Increases in the incidence of neonatal hypoglycaemia are directly related to other maternal factors, which include a poor or extreme diet, maternal diabetes, even poor economic conditions (Harris *et al.* 2012). Nevertheless, although glucose is the main oxidative fuel of the newborn brain, it is capable of utilizing the compensatory ketone bodies mobilized at birth as alternative energy sources.

Neonatal hypoglycaemia has been associated with poor neurological outcomes (Lucas *et al.* 1988). Hypoglycaemia has been widely studied since 1937; however, transient hypoglycaemia has not been that extensively documented thus far (Rozance & Hay 2010). In a recent study, Kaiser and co-workers (2015) set out to investigate if early transient hypoglycaemia, within the first 3 hours of birth, could be associated with poor learning abilities. Their experiment regulated the main outcome on the proficiency of fourth grade literacy and maths in infants who, at birth, were exposed to early transient hypoglycaemia. Interestingly, the results obtained indicated that children who were exposed to early transient hypoglycaemia compared to normoglycaemic newborns at birth were not intellectually proficient and tested with low academic scores at the age of 10 (Kaiser *et al.* 2015). These authors caution that although these results may be specific to their hospital population, their findings are noteworthy enough to serve as preliminary data for a definitive prospective trial and that further attention should be given to the current guidelines that recommend screening newborns with symptomatic hypoglycaemia or at risk of developing the condition. This recommended practice is based upon clinical experience (see 4.1.3 below) (Hawdon 2008, Cornblath *et al.* 2000).

4.1.3 The untargeted metabolomics study of cord blood from HIV-ART exposed and unexposed neonates.

In my investigation (Moutloatse *et al.* 2016) on the “Metabolomic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood” (paper presented in section 4.2), the metabolomics approach clearly highlighted the phenomenon of transient hypoglycaemia at birth, observed in the cord blood of both the HIV-ART exposed and HIV-ART unexposed neonates. Most metabolomics techniques focus on only some subsection of the human metabolome, or destroy the sample as part of the analytical procedures (e.g. in the derivatization of metabolites required for GC-MS analysis). Given the importance of this PhD project, and in view of (1) the policy of the World

Health Organization and the South African Department of Health to provide ART to all HIV-positive mothers for PMTCT, as well as (2) the recognition of the potential effect of neonatal exposure of ART on the development of children, it was our objective to cover the metabolome as extensively as possible.

Improvements in the past decade in NMR metabolomics have made it a very powerful tool for metabolomics research. Despite its low sensitivity, NMR holds many advantages over mass spectrometry (Fan & Lane 2016). The recent publication by Fan and Lane (2016) on “Applications of NMR spectroscopy to systems biochemistry”, offers a comprehensive review of the role of NMR in metabolic research. And even more recently, Dona and colleagues (2016) provide guidelines on methods for metabolite identification in biological fluids using NMR spectroscopy, providing illustrations from recently published studies. Most of these approaches featured strongly in our investigation, as I generated the NMR data in the renowned laboratory for metabolic studies using NMR technology at the Radboud University Medical Centre in Nijmegen, the Netherlands (Engelke *et al.* 2004), performed under the expert local supervision of Dr Engelke and Prof Wevers.

The outcomes of the NMR metabolomics study are presented in section 4.2, including the accompanying supplementary information, which presents additional clinical analytical and technical information on the study. The main insights that emerged from this investigation are;

1. In the cord blood analysis, I have relied mainly on chemical shifts and signal multiplicities from one-dimensional ¹H-NMR spectra, and was able to derive important spectral information for full quantitation of metabolites that are responsible for discriminating between the different neonatal groups in the study. This investigation supports the value of NMR-based cord blood analysis in identification of biological consequences of pregnancy and birth complications (this aspect is reviewed further in Chapter 7).
2. The results from the untargeted metabolomics study confirmed that metabolic changes following HIV-ART exposure are already discernible at delivery. The study also indicated that the ability of the HIV-ART exposed neonates to cope with allostasis during transitional neonatal hypoglycaemia does not seem to be significantly compromised in infants exposed to ART *in utero*. The benefit of ART for HIV-positive pregnant women, as recommended by the WHO, thus seems to be confirmed.
3. ART treatment for PMTCT nevertheless remains a matter of concern and the present study adds to the list of investigations that have indicated dysfunctional

energy metabolism in some newborns exposed to fetal ART; the short- and long-term effects of HIV-ART exposure during pregnancy on subsequent health in infancy and childhood call for continuous basic and clinical investigations.

ORIGINAL ARTICLE

Metabolic risks at birth of neonates exposed in utero to HIV- antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood

Gontse P. Moutloatse¹, Madeleine J. Bunders,^{2,3} Mari van Reenen,¹ Shayne Mason,¹ Taco W. Kuijpers,⁴ Udo F. H. Engelke,⁵ Ron A. Wevers,⁵ Carols J. Reinecke¹

¹Centre for Human Metabolomics, Faculty of Natural Sciences, North-West University (Potchefstroom Campus), Private Bag X6001, Potchefstroom, South Africa.

²Department of Experimental Medicine, Academic Medical Centre (AMC), University of Amsterdam (UvA), Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands.

³Emma Children's Hospital, AMC, University of Amsterdam. The Netherlands.

⁴Department of Pediatric Hematology, Immunology and Infectious Diseases, Emma Children's Hospital, AMC, University of Amsterdam. The Netherlands.

⁵Translational Metabolic Laboratory - 830 TML, Department Laboratory Medicine, Radboud University Medical Centre, Geert Grooteplein 10, 6525 GA, Nijmegen, The Netherlands.

Received: 2 June 2016 / Accepted: 15 September 2016
Springer Science+Business Media New York 2016

ABSTRACT

Introduction: Antiretroviral therapy (ART) for HIV-infected pregnant women is highly effective in preventing mother-to-child transmission (PMTCT) of the virus, but deleterious metabolic and mitochondrial observations in infants born to HIV-infected women treated with ART during pregnancy are periodically reported. *Objectives:* This study addresses the concern of HIV-ART-induced metabolic perturbations through a metabolomics study of cord blood collected during transitional neonatal hypoglycaemia following birth from newborns either exposed or unexposed to fetal HIV-ART. *Methods:* Proton magnetic resonance spectra from cord blood of 11 *in utero* HIV-ART-exposed and 14 unexposed newborns, as well as serum from 8 control infants, generated 114 spectral bins which were used to identify significant metabolites by means of univariate and multivariate statistical analyses. *Results* The metabolite profiles differed significantly between that from the unexposed newborns and that from infants — interpreted to characterize the state of transitional neonatal hypoglycaemia (low glucose and high lactic acid and ketone bodies). Quantitative analysis of potential ATP generation showed no meaningful difference in the global metabolite profiles of HIV-ART-exposed and unexposed neonates, but Volcano plot analysis, affirmed by odds ratios, indicated that exposure to HIV-ART affected the plasma 3-hydroxybutyric acid and hypoxanthine concentrations. *Conclusions:* The metabolite profile for transitional neonatal hypoglycaemia indicated that HIV-ART did not compromise the exposed neonates to the energy stress of allostasis experienced at birth. Increased hypoxanthine and 3-hydroxybutyric acid indicates metabolic stress at birth in some of the newborns exposed to HIV-ART and raises a concern about unrecognized prolonged allostasis with potential neurological consequences for these infants.

4.2 Introduction

The introduction of antiretroviral therapy (ART) to pregnant women to prevent mother-to-child transmission (PMTCT) of HIV, the cause of a global paediatric pandemic, has been a cornerstone in public health and one of its greatest successes (Chasela et al. 2010). Since 2001, intensified PMTCT strategies have resulted in a dramatic change in the trajectory of the AIDS pandemic and reduced the number of HIV-infected children (Doherty et al. 2013; UNAIDS 2015). Despite these impressive outcomes, ART treatment for PMTCT remains a matter of concern.

Higher mortality rates for children born to HIV-infected mothers compared to uninfected women have been reported (Newell et al. 2004; 2006; Sprecher et al. 1986). The exact underlying mechanisms are not yet clear. On balance, it seems that early *in utero* MTCT is rare (Newell 1998), and the benefit of ART to HIV-positive mothers outweighs the possible detrimental effects in the infant, with ART now being recommended on a global scale (AIDSinfo 2015; McCormack and Best 2014; Stover et al. 2011; UNAIDS 2011).

Notwithstanding the wide acceptance of this recommendation, the use of ART and especially highly active antiretroviral therapy (HAART) during pregnancy carries a caution. Cases of premature delivery and pregnancy complications have been observed following HAART (Newell and Bunders 2013). Numerous reports of deleterious metabolic and mitochondrial effects including dysfunctional intermediate metabolism in a limited number of exposed newborns, of infants and in experimental animal studies are persistently reported (Blanche et al. 1999; Barret et al. 2003 Brogly et al. 2007; Vela et al. 2008; Johnson et al. 1988). Such observations add to the concern about the clinical consequences of exposing infants to HIV-ART *in utero*. As these children increasingly form a substantial number of individuals at risk in society, there is a pressing need for longitudinal studies, from early infancy to childhood, following exposure.

Birth is the preferred point of departure for longitudinal investigations on the consequences of the offspring's exposure to ART *in utero*. During childbirth, the assured trans-placental nutrient and oxygen supply is suddenly interrupted, leading to a condition of "transitional neonatal hypoglycaemia" which is a transient phenomenon without pathological impact. Potentially severe consequences can occur due to disorders that cause severe and persistent hypoglycaemia (Stanley et al. 2015) or continued oxygen and nutrient depletion. Experiments by Kuma and co-workers with transgenic mice lacking the ability to overcome neonatal hypoglycaemia showed that these mice died within the first day of life (Kuma et al. 2004), prompting the authors to affirm that at birth the neonates "*encounter the first, and probably the most severe, period of starvation during their lifespan*" – that is, they developed a state of allostasis. Allostasis typifies perturbed homeostasis when the physiological parameters essential for life are changed by environmental conditions or life-history stages (Levine and Klionsky 2004) and depicts the ability to adapt and cope with a challenge (Singh and Cuervo 2011).

Against this background, and aware of the dissention over deleterious metabolic and mitochondrial effects on neonates and infants born to ART-exposed HIV-positive women, we formulated the following working hypothesis: ‘metabolic changes following HIV-ART exposure are expected to be discernible already at delivery’. To evaluate the working hypothesis, we defined two aims: (1) to determine the metabolite profile that characterizes transitional neonatal hypoglycaemia (Johnson *et al.* 1988)/neonatal starvation (Kuma *et al.* 2004) at birth, and (2) to record the response of neonates exposed to HIV-ART *in utero* to the perturbed homeostasis by comparing metabolite profiles of those who were either unexposed or exposed to HIV-ART. Serum samples from the infants were used to define restored homeostasis following delivery (group 1). Metabolic profiles reflecting transitional neonatal hypoglycaemia were determined using venous cord blood drawn from the umbilical cord immediately after delivery from neonates unexposed to HIV-ART (group 2), and on a possible effect due to HIV-ART, using samples from exposed newborns (group 3).

We selected proton nuclear magnetic resonance (^1H -NMR) spectroscopy as our analytical method for the present study as it permits the identification and quantification of a wide range of compounds, independent of their chemical properties. NMR is particularly amenable to detecting polar and uncharged compounds, such as sugars, amines or relatively small, volatile compounds (such as formic acid, formaldehyde, acetone, etc., which are often undetectable by LC–MS methods (Emwas *et al.* 2015). It stands out in its specificity to identify small molecule analytes in an untargeted, non-destructive, non-selective fashion from complex biological human biofluids having limited sample sizes (Dunn *et al.* 2005). The biological information from this data was determined following a combination of univariate and multivariate statistical analyses (Reinecke *et al.* 2012). A few NMR-based metabolomics studies have been reported on human pregnancy using umbilical cord blood as the biofluid (Tea *et al.* 2012; Ivorra *et al.* 2012; Walsh *et al.* 2012). However, to our knowledge this is the first metabolomics investigation to characterize the global metabolite profile of transitional neonatal hypoglycaemia and the corresponding response to fetal HIV-ART exposure. We were able to (1) define a metabolic profile prevailing during the initial phase of transitional neonatal hypoglycaemia, (2) provide metabolite information as a means to assess the benefits versus the reasons for possible concerns about newborns following HIV-ART exposure, and (3) contrive that distinct risks prevail and appropriate care should be available to neonates born to HIV-ART mothers. We believe that this study contributes to the understanding of the very beginning of the potentially life-long effect of fetal HIV-ART exposure and justifies refinement of the metabolic parameters through follow-up

investigations, including the short- and long-term effects of HIV-ART exposure in infants and children.

4.3 Methods

4.3.1 Study design

This study is divided into seven successive stages, schematically outlined in Figure 4-1. First (Figure 4-1, stage 1), serum samples obtained from infants and cord blood samples from HIV-ART unexposed neonates, collected at delivery, were analysed. Previous comparative studies on serum and plasma metabolite concentrations, using a commercial metabolomics kit in a flow-injection analysis-mass spectrometry (FIA-MS) study (Yu et al. 2011) and NMR-metabolomics (Schicho et al. 2012) indicated that either matrix should generate the same results in clinical and biological studies. One-dimensional proton NMR spectra were generated for all samples, providing the basic spectral data for this study. Subsequently the spectral bins were analysed to identify informative bins (Figure 4-1, stage 2), followed by univariate and multivariate statistical analyses to reveal bins able to differentiate between the infant and unexposed neonatal groups (Figure 4-1, stage 3). These bins were used to identify significant metabolites associated with transitional neonatal hypoglycaemia. This analysis provided baseline information for the second aim of the investigation: to evaluate the metabolic consequences for the experimental group of neonates exposed *in utero* to HIV-ART (Figure 4-1, stage 4). From these results, we used the metabolomics data to uncover energy requirements of the exposed versus the unexposed groups of neonates (Figure 4-1, stage 5) and, using this information and the original data, assessed the metabolic risk for individual neonates (Figure 4-1, stage 6) exposed to HIV-ART. Finally, we provide an interpretation of how the combined results can be used as clinical directives in dealing with neonates exposed *in utero* to HIV-ART (Figure 4-1, stage 7).

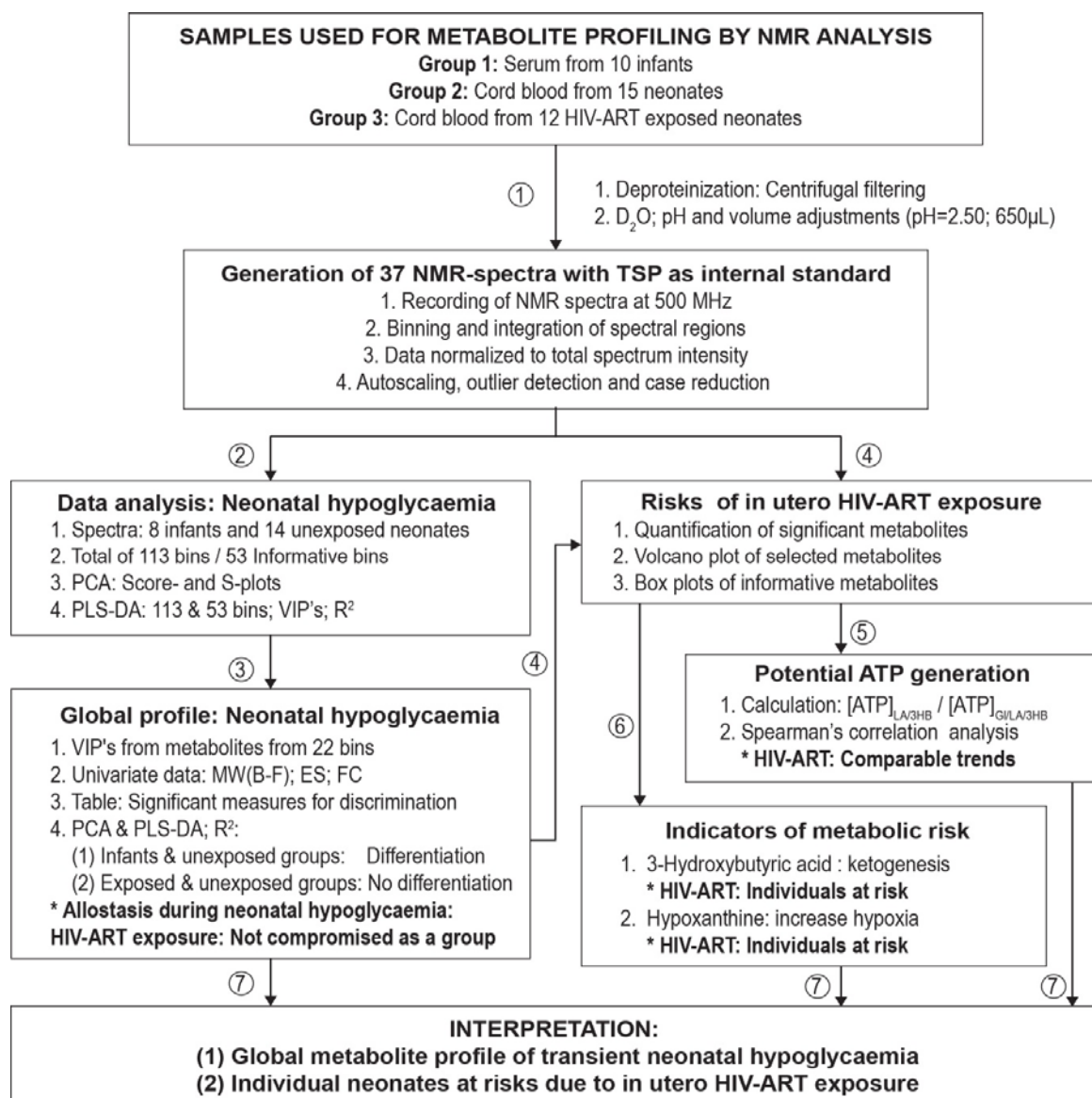


Figure 4–1: A schematic representation of the investigation workflow to generate data, to identify a metabolite profile of neonatal hypoglycaemia and to summarize the assessment of the effect of HIV-ART exposure on neonates at delivery. Circled numbers 1–7 indicate the sequence of methodological stages followed in this investigation: stage 1 represents the process towards data generation; stages 2, 3 and 7 were aimed at revealing the characteristic global metabolite profile of transient neonatal hypoglycaemia; stages 4–7 focused on identifying metabolic risks, if any, for neonates *in utero* exposed to HIV-ART. An asterisk (*) in the flow diagram indicates main conclusions formulated for the respective stages.

4.3.2 Samples and the study groups

Only biofluid samples with a volume of at least 0.8 mL (an analytical restriction) were used in the NMR analysis. HIV-related parameters such as viral load, CD4 count, antiretroviral medication and demographic data were extracted from records at the Academisch Medisch Centrum (AMC), the university hospital affiliated to the University of Amsterdam, in the Netherlands.

The characteristics of the three experimental study groups were:

(1) Infants

The complexity of obtaining matched samples from healthy infants and children in metabolomics investigations (mostly referred to as control cases) is generally recognized (Barshop 2004). In our study the control cases (group 1) consisted of infants and children aged between 8 months and 2 years, for whom serum samples (n = 10) were collected based on suspicion of an inborn error of metabolism (IEM). Following comprehensive metabolite analysis by the Potchefstroom Laboratory for Inborn Errors of Metabolism (www.pliem.co.za) in South Africa, group 1 was declared negative for IEMs. Ethical approval for the samples from the infant group was obtained from the Ethics Committees of North-West University (reference no. 02M02), including informed consent from the parents for serum analyses and eventual publication of results from anonymous cases in accordance with the ethical requirements of the diagnostic laboratory.

(2) Neonates unexposed to HIV-ART

Group 2 consisted of 15 full-term singleton neonates, hereafter referred to as the unexposed group, born vaginally after an uneventful pregnancy. The mode of delivery was comparable but not identical to the HIV-ART-exposed newborns. Inclusion criteria for this control group were healthy singleton babies with a gestational age of ≥ 37 weeks with a weight appropriate for gestational age and an uneventful pregnancy and delivery. Exclusion criteria for the control group were neonatal low birth weight, perinatal asphyxia, HIV-infected mothers and any clinical or genetic abnormality in the neonates.

Samples of venous cord blood were obtained from the clamped umbilical cord immediately after delivery, collected in chilled heparin containing vacutainers, followed by centrifugation ($5000 \times g$, 10 min, 4°C). Aliquots of cord plasma were stored at -20°C prior to analysis and transported to the NMR facility at the Translational Metabolic Laboratory at the Radboud University Medical Centre in Nijmegen for NMR analysis. Cord blood from the ART-exposed and unexposed neonates was collected in the same way.

(3) Neonates exposed to HIV-ART

Table A1-1 (Annexure 1) summarizes the clinical parameters for the mothers of the ART-exposed neonates ($n = 12$), hereafter referred to as the exposed group. The mean age of these HIV-infected women was 32 ranging from 21 to 40 years of age at time of delivery. There was a mean nadir CD4 cell count of 235 (range: 40–920), and as part of the clinical follow-up during the pregnancy, CD4+ T-cell count and plasma HIV-RNA were measured during the first, second and third trimesters of pregnancy. Almost all women had been diagnosed with HIV before this pregnancy and received combination-ART (cART) from first trimester. The cART of the mothers was not identical, but reflected the real world situation of ART adhered to by HIV-positive mothers according to prescription by their clinicians. According to European guidelines, the cART profile consisted of a nucleoside reverse transcriptase inhibitor (NRTI) backbone combined with either a protease inhibitor or the non-nucleoside reverse transcriptase inhibitor (NNRTI) nevirapine. The mean gestational age was 39.1 (38.3–41.1) weeks. Modes of delivery were vaginally ($n = 7$), vaginally with induction ($n = 1$), and caesarean section ($n = 4$).

Table 4-1 summarizes the clinical parameters of the exposed neonates (group 3). The mean birth weight of the newborns was 3449 (2900–4620) grams. Inclusion criteria for the case study group were gestational age ≥ 35 weeks and birth weight > 2500 g (appropriate for gestational age). The exclusion criteria were HIV-exposed neonates who tested positive to any other bacterial or viral infections detected. Venous umbilical cord blood samples were obtained within 10 minutes of delivery. Ethical approval for the study and for the use of the samples from the two neonatal groups was obtained from the Ethics Committee of the University of Amsterdam.

Table 4–1: Clinical data details of the neonates *in utero* exposed to HIV-ART.

Gender (M:male; F:female)	Ethnicity	Gestational age (weeks)	Mode of delivery*	Birth weight (g)	Apgar (1,5,10 min)	score
M	Caucasian	39.4	Vaginal (Normal)	4620	8	10
M	Black	38.5	Caesarean (Normal)	4090	9	10
M	Black	40.2	Vaginal (Normal)	3170	8	10
M	Black	39.2	Vaginal (Normal)	3560	6	8 10
M	Black	38.5	Caesarean (Normal)	3440	7	9 10
F	Black	38.3	Vaginal (Normal)	3035	10	
M	Black	40.1	Caesarean (Normal)	3465	10	10
F	Black	39.3	Vaginal (Normal)	2900	9	10
F	Black	38.3	Vaginal (Normal)	3185	9	10
M	Black	37.6	Vaginal (Induced)	3850	10	10
F	Black	39.1	Vaginal (Normal)	2565	4	7 8
M	Caucasian	41.1	Caesarean (Normal)	3506	8	10 10

*: Normal indicated mothers who gave birth without prior induction.

4.3.3 Sample preparation and metabolomics analysis

Sample preparation was conducted according to the methodological design (Figure 4-1: stage 1) and by using the standard operational procedure established by the Nijmegen Translational Metabolic Laboratory (Engelke *et al.* 2006, 2008; Wevers *et al.* 1995), as described in detail in the Supplementary Information (SI; Annexure A1). Plasma and serum samples contain large molecular weight proteins and lipoproteins, which affect the NMR spectroscopic identification of smaller molecular weight metabolites, which could be avoided by deproteinization through ultrafiltration. Once deproteinized, 700 μ L of the supernatant (filtrate) was transferred into a test tube and diluted with 20 μ L internal standard solution [freshly prepared deuterium oxide (D_2O) solution containing 20.2 mmol/L 3-trimethylsilyl-tetradecuteriopropionic acid (TSP, sodium salt; Merck, Darmstadt, Germany)] as the chemical shift reference (δ = 0.00 ppm) and as a deuterium lock signal. The pH of the samples was then altered to 2.50 ± 0.05 . A final volume of 650 μ L of the prepared samples was placed in 5-mm high quality NMR tubes (Wilmad Royal Imperial; Aldrich) for analysis.

High-resolution NMR spectra of all samples were recorded on a 500 MHz ^1H resonance frequency spectrometer (AMX-500; Bruker Analytische Messtechnik, Karlsruhe, Germany). The samples were loaded in a sample changer in a random order to avoid any confounding machine variation. All spectra were recorded at a

temperature of 298 K. NMR data pre-processing was performed using Bruker Topspin (V3.1) and Bruker AMIX (V3.9.12) was used for binning and quantification (Figure 4-1: phase 2). Division into bins across the spectral regions was performed automatically in AMIX where variable-sized bucketing (van den Berg *et al.* 2006; Cui and Churchill 2003; De Meyer *et al.* 2008) was used to create a data matrix of 114 bins/variables (Tables A1-2, A1-3 and A1-4).

Spectral regions devoid of peaks (noise) and the region containing water resonances were excluded from the data. All data were normalized in AMIX by dividing each integrated segment by the total spectrum intensity (minus the excluded regions), which compensated for the differences in overall concentration between individual plasma samples. Assignment of potential metabolites to each bin resulted in 54 bins with a distinct metabolite. Pure compound NMR spectral libraries at pH 2.5 were used to identify metabolites as described by Wevers *et al.* (1994). The remaining bins contained substances and features which could not be annotated unequivocally as a distinct metabolite. Statistical analysis revealed that the 54 bins contained meaningful information on transient neonatal hypoglycaemia and on fetal HIV-ART exposure, which was not observed for the excluded bins (A1-2, A1-3).

4.4 Statistical analysis

Statistical analyses were performed using Matlab (MATLAB with Statistics and PLS Toolbox Release (2012b), The MathWorks, Inc., Natick, MA, USA), SAS (SAS Institute Inc. (2015). The SAS System for Windows Release 9.3 TS Level 1M0 Copyright© by SAS Institute Inc., Cary, NC, USA) and SPSS (SPSS Inc. (2015). IBM SPSS Statistics Version 22, Release 22.0.0, ©IBM Corporation and its licensors. <http://www-01.ibm.com/software/analytics/spss/>).

The data presented for the statistical analysis comprised 12 HIV-ART-exposed cases (cord blood samples), 15 unexposed cases (cord blood samples) and 10 infant cases (serum samples). Data pre-treatment included data transformation, centring and outlier detection (Figure 4-1: stage 2). Outliers were identified using a 95 % confidence region in a Hotelling's T² analysis in conjunction with a principal component analysis (PCA) score plot (van den Berg *et al.* 2006; Eriksson *et al.* 2006), and confirmed as outliers based on clinical and analytical information. Following the removal of outliers, the final three groups consisted of information from 11 HIV, 14 control and 8 infant cases. Only one bin, identified to represent inosine (6.09 ppm), was removed as it occurred in only two cases (50 % rule). The 113 bins served to identify differences between the groups

through univariate as well as multivariate analysis using various supervised and unsupervised techniques.

Pre-processing was done at a univariate level using auto-scaling to place all variables (bins) on equal footing (van den Berg *et al.* 2006). Though auto-scaling may inflate measurement error (van den Berg *et al.* 2006), this is of lesser concern for NMR-generated data. Univariate analysis of the three groups without outliers included fold changes (FC), Mann–Whitney (MW) *p*-values and effect sizes (ES). These were calculated for pair-wise comparisons between groups on the untransformed, unscaled data (Cui & Churchill 2003). Finally, unsupervised PCA was performed on the scaled data to identify natural separation between groups. PCA was conducted using the first three components respectively, with 90 % confidence ellipses for each group. Partial least squares discriminant analysis (PLS-DA) was the only supervised multivariate analysis performed on the scaled data. Scores for variables important in discrimination (VIP) were used in combination with univariate statistics to select a short list of significant variables for quantification and biological interpretation, following the sequence shown in Figure 4-1. Quantified variables (metabolite concentrations) were used to construct logistic regression models to evaluate odds ratios (ORs).

4.5 Results and Discussion

4.5.1 Data generation and case selection

A representative scaled NMR spectrum from each of the three experimental groups is given in Figure. A-1 to illustrate some of the discernible qualitative NMR differences observed in the selected examples from the infant, unexposed and exposed neonatal groups (Figure A-1a–A1-1c; procedure shown in Figure 4-1: stage 2). Calculating the peak areas within specified segments of a spectrum (binning) (De Meyer *et al.* 2008) allowed comparison of NMR data measured for different samples, including identification of metabolites, other variables or unknowns/features associated with each bin. Following compilation of the spectral data, case reduction was applied according to the protocol outlined in Figure 4-1. Four outliers (two from the infant group and one each from the exposed and unexposed neonatal groups) were identified using a 95 % confidence region in a Hotelling's T^2 test in conjunction with the respective PCA score plots (Figure. A1-2a–2f, justified from clinical and analytical information as indicated in the SI; Annexure1).

4.5.2 Metabolite profile of cord blood of neonates unexposed to HIV-ART at birth

The unsupervised PCA based on data from the infant and unexposed neonatal groups from the 113 normalized bins (Data provided in Tables A1-2 – A1-4 for all three experimental groups and Figure A1-3 for the PCA of all groups), is illustrated in Figure 4-2. The PCA score plot (Figure 4-2a) indicates a complete separation between the infant controls and the unexposed neonates, with the first principal component (PC1) accounting for 31 %, and PC2 for 15 %, of the variability in the data. The PCA loadings S-plot (Figure 4-2g) allows easy visualization of the main trends in the data. The two bins at the lower left end correspond to lactic acid (elevated in the unexposed neonates), whereas seven bins at the upper right end are associated with glucose and one with carnitine (both metabolites decreased in the unexposed neonates). Univariate analysis (Table A1-5) substantiated the significant differences between the groups in terms of the spectral data for these metabolites, by having $p < 0.05$ for the Mann–Whitney tests and $FC > 2.0$ or $FC \leq -2.0$ for the respective fold changes (+: upregulated; –: down-regulated for the neonates relative to the infants).

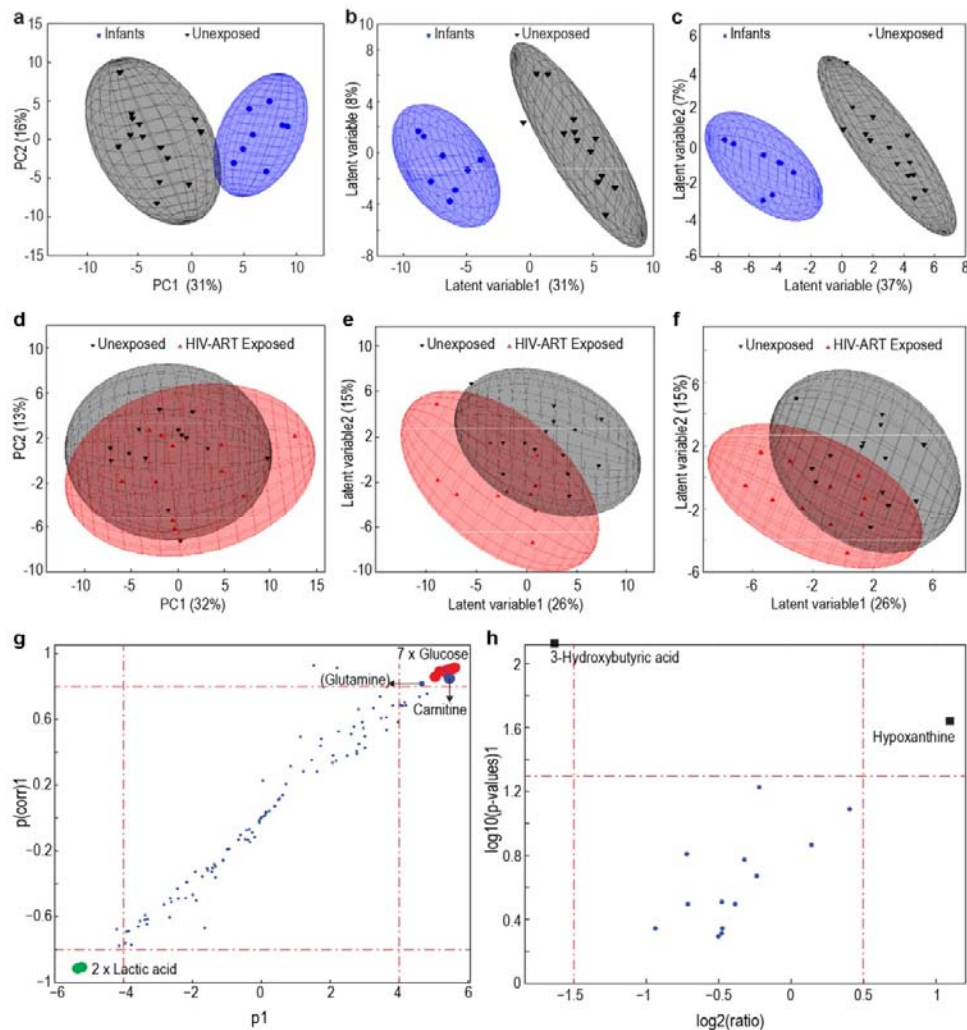


Figure 4–2: Statistical analyses to identify metabolites being important during transitional neonatal hypoglycemia in neonates exposed or not exposed to HIV-ART *in utero*. (a) PCA score plots of infant and unexposed neonatal groups based on the normalized data from the 113 bins. (b) PLS-DA score plots of infant and unexposed neonatal groups based on the 113 bins. (c) PLS-DA score plots of infant and unexposed newborns based on the normalized data from the 53 bins. (d) PCA score plots of unexposed and exposed newborns based on the 113 bins. (e) PLS-DA score plots of exposed and unexposed neonatal groups based on the 113 bins. (f): PLS-DA score plots of unexposed and exposed newborns based on the normalized data from the 53 bins (g) PCA S-plot for the infant and unexposed neonatal groups, using spectral data from the 113 bins. (h) Volcano plot of the 15 quantified metabolites from the neonatal groups exposed or unexposed to HIV-ART.

The supervised PLS-DA (based on the data from the 113 normalized bins) is presented in Figure 4-2b, indicating a complete separation between the infant controls and unexposed neonatal groups. Annotation of a single metabolite could be made unequivocally for 53 of the 113 bins (Figure 4-1: stage 3). Complete separation was again observed through the PLS-DA (Figure 4-2b), from which 22 bins were identified to be the most important for the differentiation between the two experimental groups ($VIP > 1.0$). The 22 bins could be uniquely associated with 11 metabolites — spectral values summarized in Table A1-4. Evaluation of the PLS-DA models applied to spectral data from the 113, 53 and 22 bins, respectively, was performed using the goodness-of-fit parameter R^2 (variation in class membership explained by the model). The result was good model fit values in all three cases ($R^2 = 0.95, 0.94$ and 0.89 , respectively). The cross-validated fit statistics (Q^2) for the first latent variable were above 0.7 for all three comparisons.

The key finding from the statistical analysis of the spectral data is that these 11 metabolites (Table A1-5) provide qualitative information on the hypoxic and hypoglycaemic state prevailing at birth (Figure 4-1: stage 4), with carnitine, glucose and lactic acid being the three most significant markers ($VIP \geq 2.0$; $p = 0.0001$ and $FC \leq +2.0 / \geq -2.0$). Decreased free carnitine in the neonates ($p = 0.0001$; $FC = -2.46$) could reflect a reduction in intrauterine maternal-fetal transfer in late gestation (Novak *et al.* 1981; Glasgow *et al.* 1983). Depletion of glycogen is implicated in the significant low levels of glucose in the neonates ($p = 0.0001$; $FC = -2.67$) presenting with hypoglycaemia as the most common manifestation at, and following, birth. Increased levels of lactic acid ($p = 0.0001$; $FC = 3.14$) and ketone bodies (3-hydroxybutyric acid, $p = 0.0015$; $FC = 2.33$) indicate their participation as additional key energy resources for the critical energy dependency of the newborn brain. All amino acids with $VIP \geq 1.0$ (Table A1-4: phenylalanine, glutamine, tyrosine and proline) were decreased, which might indicate the start of gluconeogenesis or other energy generation reactions from amino acids. Data on methionine, a glucogenic amino acid and the only elevated metabolite reported to be unique to neonatal hypoxic ischemic encephalopathy (HIE) versus matched controls (Reinke *et al.* 2013), is not included in Table A1-5, because of limitations in monitoring it in NMR spectra (discussed in A1-6). Finally, creatine ($p = 0.0043$; $FC = -1.37$), which is the main endogenous source of creatinine ($p = 0.0009$; $FC = 1.85$), is a metabolite that is well known to increase in the neonatal state, having a maternal origin *in utero* and at birth (Novak *et al.* 1981).

4.5.3 Metabolite profile of cord blood of neonates exposed to HIV-ART at birth

The same analytical procedure described above was applied to the groups of neonates, both exposed and unexposed (the latter acted as controls) to HIV-ART *in utero*, to investigate the potential effect at delivery of fetal HIV-ART exposure (Figure 4-1: stage 4). The comparable results are shown in Figure 4-3. First, a PCA was performed on the spectral data from the 113 bins (Tables A1-2 and A1-3). No sample clustering or natural differentiation of the exposed and unexposed groups was observed (Figure 4-2d), in contrast to the results obtained for the infants and unexposed neonates (Figure 4-2a). Second, supervised PLS-DA for the spectral data from the 113, 53 (Figure 4-2e and 4-2f) and 22 bins also produced no separation between the two groups. The respective goodness-of-fit values were $R^2 = 0.57$, 0.57 and 0.49 , respectively. Taken together, the key conclusion based on the results from Figure 4-2d – 4-2f is that the global serum metabolite profiles of the unexposed and exposed neonatal groups were closely similar.

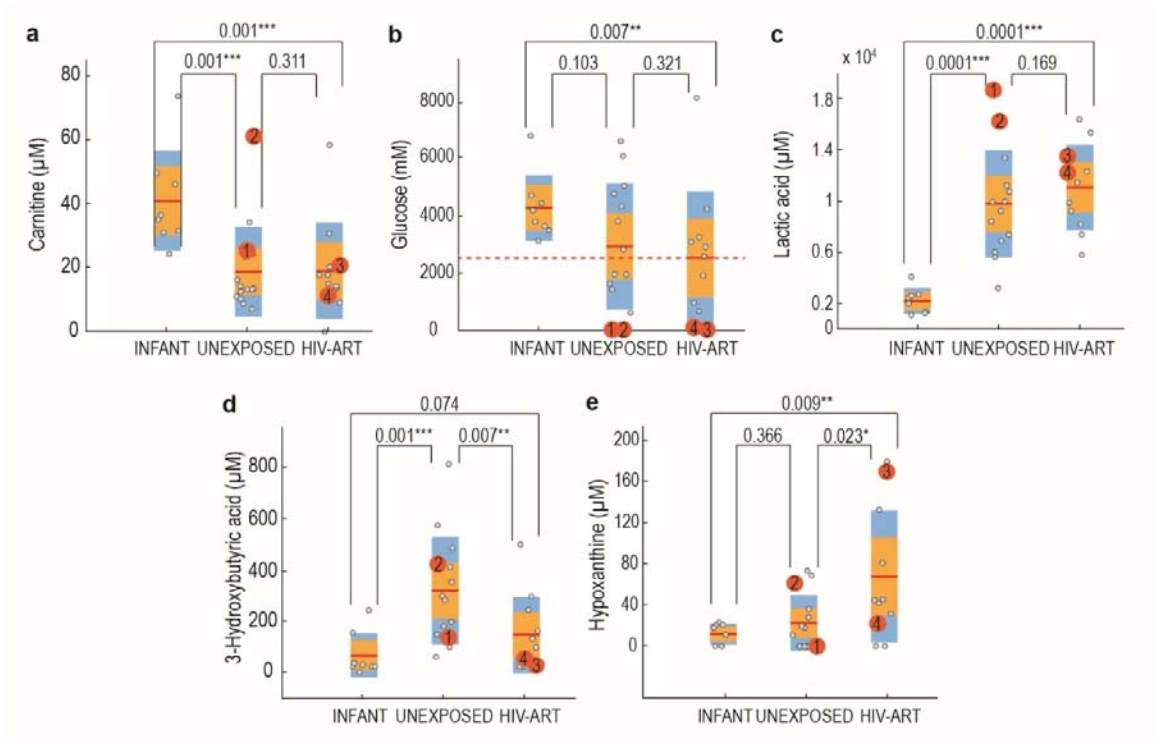


Figure 4–3: Graphs showing important metabolites indicating transient neonatal hypoglycaemia. Indicated in the figure are: Infants relative to unexposed neonates (3a–3c) and HIV-ART exposure cases relative to infants or unexposed cases (3d–3e). Values for all individual cases are shown as dots, while the squared area represents the 95 % confidence interval (red) and 1 standard deviation (blue) of the mean (red line). The statistical significance of differences (i.e. *p*-values) is indicated for every group relative to the other, in the upper section of each set of graphs (*: *p* < 0.05; **: *p* < 0.01; ***: *p* < 0.001). The threshold of fasting hypoglycaemia (2.5 mM) is indicated by the red dotted line in (b). The profiles of four newborns with extreme hypoglycemia are highlighted, shown in orange and numbered 1 and 2 for the two unexposed and 3 and 4 for the two HIV-ART-exposed cases. The values for these four cases are likewise indicated for the remaining metabolites (3d and 3e), as well as in Figure 4-4.

Note: Quantification of these 5 metabolites was done on the following spectral data bins: (a) carnitine (3.22 ppm); (b) glucose (5.22 ppm); (c) lactic acid (1.41 ppm); (d) 3-hydroxybutyric acid (1.23 ppm); (e) hypoxanthine (8.42 ppm).

As the multivariate analyses did not reveal a noteworthy effect of HIV-ART exposure, we subsequently performed univariate analyses (Figure 4-1: stage 6) in an alternative attempt to uncover effects of exposure to HIV-ART, if any. We scrutinized the data from the 53 informative bins, revealing that spectral information in 4 bins, linked to 3-hydroxybutyric acid, hypoxanthine, myoinositol and dimethyl sulfone (DMSO₂), showed a value for ES > 0.3 and a Mann–Whitney $p < 0.05$. The concentrations of these 4 metabolites and of the 11 important metabolites previously identified (Table A1-5) were calculated for cord blood from the exposed and unexposed neonates and are summarized in Table 4-2. All 15 metabolites exhibit FC $\geq +1.0$ / ≤ -1.0 , indicating that they could potentially contribute to the differentiation of the experimental groups from which they were derived. A Volcano plot identified two metabolites that indicated an effect of HIV-ART exposure *in utero* (Figure 4-2h). The two metabolites were (1) decreased 3-hydroxybutyric acid ($p = 0.007$; ES = 0.49; FC = -2.19) and (2) increased hypoxanthine ($p = 0.023$; ES = 0.40; FC = 3.02).

Table 4–2: Quantitative data on the HIV-ART-exposed and unexposed neonates for important metabolites identified from the spectra corresponding to the two experimental groups. The chemical shift for the specific bins used for the quantification is shown in brackets after the name of the metabolite.

Metabolite (ppm bin quantified)	Mann–Whitney <i>p</i> -value	Mann–Whitney effect size	Fold change	Mean (μM) Exposed	Mean (μM) Unexposed	SD Exposed	SD Unexposed
3-Hydroxybutyric acid (1.23)	0.007	0.49	–2.19	145	318	149	210
Hypoxanthine (8.42)	0.023	0.40	3.02	69	23	65	27
Histidine (8.65)	0.059	0.31	1.21	68	56	18	14
Myoinositol (4.05)	0.081	0.28	1.87	450	241	618	305
Formic acid (8.24)	0.137	0.22	1.56	102	66	69	33
Creatinine (4.29)	0.156	0.20	–1.17	44	51	14	18
Lactic acid (1.41)	0.169	0.19	1.13	11079	9797	3316	4176
Creatine (3.05)	0.214	0.16	1.20	77	64	33	18
Carnitine (3.22)	0.311	0.10	1.01	19	19	15	14
Glucose (5.22)	0.321	0.09	–1.16	2559	2967	2345	2217
Phenylalanine (7.32)	0.321	0.09	1.08	62	57	19	22
DMSO ₂ (3.14)	0.456	0.02	–1.35	12	16	12	23
Glutamine (2.47)	0.456	0.02	1.02	404	397	210	225
Tyrosine (6.89)	0.489	0.01	1.01	67	66	21	23
Leucine (0.95)	0.511	0.01	–1.00	130	130	42	37

SD, standard deviation

4.5.3.1 Metabolic risks seen at birth for neonates exposed to HIV-ART in utero

The profiles of the three most important indicators of allostasis associated with transitional neonatal hypoglycaemia are shown for all individual cases studied as box plots in Figure 4-3a – 4-3c. The main statistically significant changes during neonatal hypoglycaemia are decreased carnitine (Figure 4-3a), reduced glucose (Figure 4-3b) and elevated lactic acid (Figure 4-3c) concentrations. Although lower, the *p*-values for carnitine, glucose and lactic acid did not differ significantly between the two groups of neonates (*p* = 0.311, 0.321 and 0.169, respectively). This statistical information underscores our observations from the multivariate analyses that HIV-ART-exposed neonates as a group did not seem to be compromised by their *in utero* exposure in responding to the energy stress presented by transitional neonatal hypoglycaemia.

This statistical argument does not, however, suffice in dealing with potential risks for individual HIV-ART-exposed neonates. First, cord blood glucose values below 2.5 mmol/L (a threshold value for fasting hypoglycaemia) were observed for seven out of nine exposed newborns, reaching zero and near-zero values in a few cases. However, a note of caution is required as eight of 14 unexposed cases likewise showed such low glucose concentrations. In addition to the well-established physiological reduction in plasma glucose following birth (Stanley *et al.* 2015), decreased glucose may be an artefact due to continued glycolysis in heparinized blood samples if the samples are not sufficiently cooled and centrifuged soon after collection (Bruns *et al.* 2009). Each molecule of glucose degraded during *in vitro* glycolysis produces two molecules of lactic acid if no fluoride is present in the heparin-containing tubes used in sampling (Mikesh & Bruns 2008), as was the case in our study. Nonetheless, sufficient glucose provision to the fetal and neonatal brain is critical (Stanley *et al.* 2015) and clinical interventions are indicated if the blood glucose concentration drops below 2.5 mM (Cornblath *et al.* 2000).

To investigate further the supply of energy fuels (glucose, lactic acid and 3-hydroxybutyric acid) for the neonatal metabolic adaptation in transitional neonatal hypoglycaemia, we expanded our study on individual risks by comparing two supposed scenarios on *in vivo* energy provision at birth: (1) potential ATP generation from lactic acid and 3-hydroxybutyric acid (designated $[ATP]_{LA/3HB}$ — see formula in SI Table A1-6), and (2) potential ATP from glucose, lactic and 3-hydroxybutyric acids (designated $[ATP]_{GI/LA/3HB}$). The potential ATP generated in both scenarios for unexposed neonates as a function of cord blood glucose concentration below 3.0 mM is shown in Figure 4-4a, indicating an inverse relation between ATP potential and glucose concentration (for

reference purposes, the samples with extremely low glucose mentioned above are shown as orange dots in Figure 4-3 and 4-4). For $[ATP]_{LA/3HB}$ Spearman's correlation coefficient was $r = -0.731$ ($p = 0.04$) with $R^2 = 0.793$ for the regression line, and the theoretical potential ATP available at zero glucose was 251 mM — this shows that the theoretical absence of glucose in all these cases is fully compensated for by the alternative metabolic fuel provided by lactic acid, supplemented by 3-hydroxybutyric acid. The correlation between low glucose (<2.5 mM) and $[ATP]_{GI/LA/3HB}$ is expressed by Spearman's correlation coefficient $r = -0.539$ ($p = 0.168$) with $R^2 = 0.352$ for the regression line. The theoretical potential ATP available at zero glucose was the same (251 mM) as for $[ATP]_{LA/3HB}$.

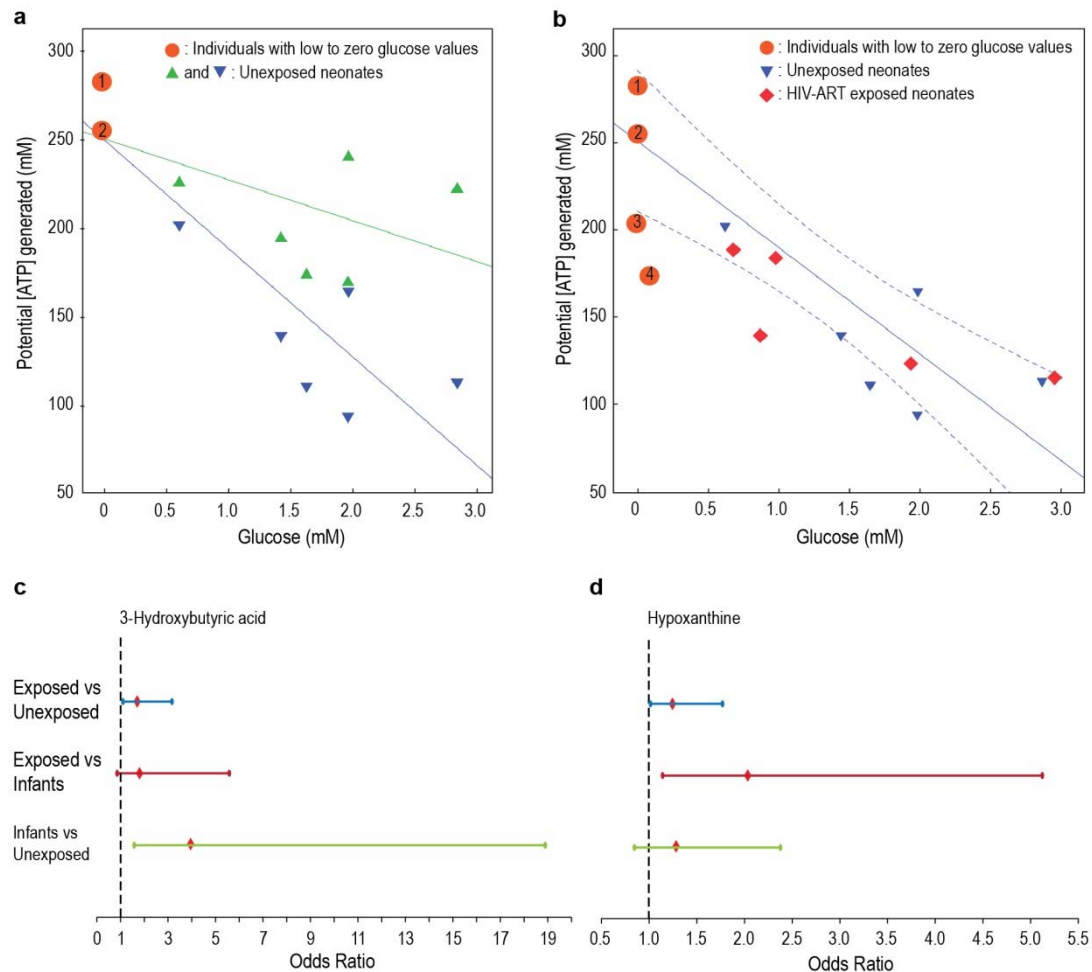


Figure 4–4: Regression lines of potential ATP generation and OR plots of important informative metabolites for all individual case studies. (a) Data on low (0–2.5 mM) cord blood glucose for unexposed neonates according to $[ATP]_{LA/3HB}$ (lower curve – blue) and $[ATP]_{GI/LA/3HB}$ (upper curve – green). (b) Individual values for the HIV-ART-exposed neonates (red) to compare ATP generation according to $[ATP]_{LA/3HB}$ against low cord blood glucose for the unexposed neonates (solid line – blue) including its 90 % confidence intervals (dotted lines – blue). (c) Odds ratio plots with 95 % confidence intervals for 3-hydroxybutyric acid. (d) Odds ratio plots with 95 % confidence intervals for hypoxanthine. The estimated OR is indicated by the red diamond on the horizontal line and the vertical dashed line indicates significance if CI does not contain 1.0.

Figure 4-4b indicates the consequences of HIV-ART exposure for potential ATP generation. We compare exposed and unexposed newborns by plotting the $[ATP]_{LA/3HB}$ values for the former relative to the $[ATP]_{LA/3HB}$ regression line of the latter ($R^2 = 0.715$), also showing the 90 % confidence bounds. The calculated ATP generation for three of seven exposed and two of eight unexposed hypoglycaemic neonates fell outside the 90 % confidence bounds (Figure 4-4b). Our key conclusion is that the metabolic response towards hypoglycaemia (glucose <2.5 mM) at birth for the neonates exposed to HIV-ART shows a trend of energy compensation comparable with that of the unexposed. The response to allostasis at birth did not seem to be compromised for the exposed newborns in our experimental group, substantiating the results and conclusions derived from Figure 4-2d – 4-2f.

4.5.3.2 Significant discriminating metabolites on in utero HIV exposure

Two metabolites (3-hydroxybutyric acid and hypoxanthine), observed to differentiate between HIV-ART-unexposed and exposed neonates (Figure 4-2h), indicated significant differences for individual cases (Figure 4-3d – 4-3e). In the transition at birth, the liver switches from carbohydrate utilization to fatty acid oxidation and ketone body production (McGarry & Foster 1980). Ketone bodies are especially important in maintaining cerebral energy supplies. The reduced 3-hydroxybutyric acid concentrations, as observed in the neonates exposed to HIV-ART (Figure 4-3d) and reported for preterm neonates (Hawdon *et al.* 1992), indicate a combined failure of lipolysis and ketogenesis, underscoring comparable risks for the preterm and some ART-exposed neonates who demonstrated extremely low 3-hydroxybutyric acid values. Further, newborns with plasma hypoxanthine levels above $25 \mu\text{mol/L}$ can manifest a higher risk of dying (Saugstad 1975). Cord blood levels of hypoxanthine greater than $25 \mu\text{mol/L}$ demonstrated this metabolic risk in seven of the 11 HIV-ART-exposed neonates in this study (Figure 4-3e) as hypoxanthine has been proposed as a specific indicator of hypoxia in the newborn (Saugstad 1975).

Finally, odds ratio estimates (ORs) with 95 % confidence intervals (CIs) were calculated for 3-hydroxybutyric acid and hypoxanthine for the three experimental groups, using logistic regression models. Figure 4-4 shows the relationship for 3-hydroxybutyric acid (Figure 4-4c) and hypoxanthine (Figure 4-4d) and HIV-ART exposure, with a unit increase corresponding to 85 and $10 \mu\text{mol/L}$, respectively. OR values less than 1 were inverted to ease interpretation (Osborne 2006; McHugh 2009). The OR for 3-hydroxybutyric acid was 1.69 (95 % CI: 1.08–3.15), in agreement with the observation of increased and beneficial ketogenesis during transitional neonatal

hypoglycaemia. This value indicates impaired ketogenesis associated with HIV-ART exposure, however — the OR values for exposed newborns vs. infants and exposed newborns vs. unexposed neonates were 1.77 (95 % CI: 0.85–5.56) and 3.94 (95 % CI: 1.56–18.9)], respectively. The OR for hypoxanthine does not appear to be significant for unexposed newborns vs. infants [OR: 1.29 (95 % CI: 0.85–2.36)]. However, an infant is twice as likely [(OR: 2.01 (95 % CI: 1.14–5.130)] to present clinically as one exposed for every 10 $\mu\text{mol/L}$ increase in hypoxanthine in cord blood plasma. We postulate that an increase of 20 $\mu\text{mol/L}$ hypoxanthine in plasma from cord blood for newborns known to be exposed to HIV-ART presents a risk of potential long-term consequences of exposure to ART exposure *in utero* and justifies the monitoring of future development.

4.6 Concluding remarks

The results presented in Table 4-3 confirmed our working hypothesis that metabolic changes following HIV-ART exposure are already discernible at delivery. The findings shown in Figure 4-3d, 4-3e, 4-4c and 4-4d as well as those of other investigations (Blanche *et al.* 1999; Barret *et al.* 2003; Brogly *et al.* 2007; Vela *et al.* 2008; Johnson *et al.* 1988) continue to raise concerns regarding the safety of fetal ART exposure, despite observations that newborns exposed to ART treatment appear to be healthy at birth. We are aware that our findings have the intrinsic limitation of relatively small experimental groups, not unlike most metabolomic investigations on complex perturbations that rely on human samples. The outcomes revealed here nevertheless justify a more comprehensive investigation of the metabolic profile of newborns exposed to HIV-ART. This study indicated that the ability to cope with allostasis during transitional neonatal hypoglycaemia does not seem to be significantly compromised in infants exposed to HIV-ART *in utero*. Although, the outstanding benefits of ART for HIV-positive pregnant women and their offspring are clear (as recommended by the WHO), ART treatment for PMTCT remains a matter of concern as the present study adds to the list of investigations that have indicated dysfunctional energy metabolism in some newborns exposed to fetal ART. The short- and long-term effects of HIV-ART exposure during pregnancy on infancy and childhood call for continuous basic and clinical investigations. Meticulous clinical monitoring is recommended for newborns whose HIV-positive mothers may have been exposed to ART during pregnancy. In addition, as these children increasingly form a substantial number of individuals at risk in society, there is a pressing need for longitudinal studies, from early infancy to childhood, following exposure. The results presented here provided information for

such an investigation, as birth is the preferred point of departure for longitudinal investigations on the consequences of the offspring's exposure to ART in utero.

Abbreviations AMC: Academisch Medisch Centrum; ART: Antiretroviral therapy; cART: Combination antiretroviral therapy; CI: Confidence interval; DMSO₂: Dimethyl sulfone; D₂O: Deuterium oxide; ES: Effect size; FC: Fold change; HAART: Highly active antiretroviral therapy; HIV: Human immunodeficiency virus; ¹H-NMR: Proton nuclear magnetic resonance; IEM: Inborn errors of metabolism; MTCT: Mother-to-child-transmission; MW: Mann–Whitney; NMR: Nuclear magnetic resonance; NNRTI: Non-nucleoside reverse transcriptase inhibitor; NRTI: Nucleoside reverse transcriptase inhibitor; OR: Odds ratio; PCA: Principal component analysis; PC: Principal component; PLS-DA: Partial least squares discriminant analysis; PMTCT: Preventative mother-to-child transmission; SI: Supplementary information; TSP: Trimethylsilyl-propionic acid; VIP: variables important in projection; WHO: World Health Organization.

Author's contributions This investigation required a multidisciplinary approach and the inputs of all authors were essential to produce the final manuscript. CR and MJB defined the aim of the study, developed the experimental design and obtained ethical approval for the study. MJB and TWK provided the neonatal samples and clinical description on the neonates and their mothers. CR provided it for the infant controls and was responsible for coordination and integration of inputs from the authors. GM conducted all experimental analyses and compiled the clinical information provided by MJB. GM, SM and UE gave guidance and assessment on ¹H-NMR data generation and spectral analyses. RW was responsible for critical evaluation of the analytical aspects of the clinical chemistry data and of their interpretation. MvR performed all the statistical analyses. All authors have given final approval of the version of the manuscript as submitted and agreed to the accountability requirements.

Acknowledgements Research funding for this project was provided by the Technological Innovation Agency (TIA) of the Department of Science and Technology of South Africa. GM received a post-graduate bursary from the National Research Foundation (NRF) of South Africa and a SAVUSA-SKILL stipend from the Embassy of The Netherlands in Pretoria, enabling her to pursue the NMR analyses at the Radboud University Medical Centre in Nijmegen. The views and findings of this investigation are those of the authors and do not reflect or represent any policies of the NRF, TIA or

SAVUSA-SKILL funders. Further support was obtained from by the Netherlands Organisation for Scientific Research Agiko stipendium (92003417) and the Dutch Aids Foundation (2006015). The funders were not involved in the writing of the manuscript, neither in its editing nor in the decision for it to be distributed for publication. We would like to thank K. Boer, M. Godfried, J. Nellen, E. van Leeuwen and H.J. Scherpbier for clinical care of the pregnant women and their children. Furthermore, we thank J.L. van Hamme for sample processing.

Compliance with ethical standards

Competing interests The Authors declare that there is no conflict of interest.

Ethical approval

1. Neonates: Ethical approval for the study and for the use of the samples from the two neonatal groups was obtained from the Ethics Committee of the University of Amsterdam.

2. Infants: Ethical approval for the samples from the infant group was obtained from the Ethics Committees of North-West University (reference no. 02M02), including informed consent from the parents for serum analyses and eventual publication of results from anonymous cases in accordance with the ethical requirements of the diagnostic laboratory (Potchefstroom Laboratory for Inborn Errors of Metabolism - www.pliem.co.za).

Supplementary information: see Annexure 1.

REFERENCES

SECTION 4.1

Cornblath, M. (1997). Neonatal hypoglycemia 30 years later: does it injure the brain? Historical summary and present challenges. *Acta Paediatrica Japonica; Overseas edition*, 39, S7–S11.

Cornblath, M., Hawdon, J. M., Williams, A. F., Aynsley-Green, A., Ward-Platt, M. P., Schwartz, R., & Kalhan, S. C. (2000). Controversies regarding definition of neonatal hypoglycemia: suggested operational thresholds. *Pediatrics*, 105(5), 1141–1145.

Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., & Everett, J. R. (2016). A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Computational and Structural Biotechnology Journal*, 14, 135–153.

Engelke, U. F., Liebrand-van Sambeek, M. L., De Jong, J. G., Leroy, J. G., Morava, É., Smeitink, J. A., & Wevers, R. A. (2004). N-acetylated metabolites in urine: proton nuclear magnetic resonance spectroscopic study on patients with inborn errors of metabolism. *Clinical Chemistry*, 50(1), 58–66.

Fan, T. W. M., & Lane, A. N. (2016). Applications of NMR spectroscopy to systems biochemistry. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 92, 18–53.

Harris, D. L., Weston, P. J., & Harding, J. E. (2012). Incidence of neonatal hypoglycemia in babies identified as at risk. *The Journal of Pediatrics*, 161(5), 787–791.

Hawdon, J. M. (2008). Investigation and management of impaired metabolic adaptation presenting as neonatal hypoglycaemia. *Paediatrics and Child Health*, 18(4), 161–165.

Hawdon, J. M., Platt, M. W., & Aynsley-Green, A. (1992). Patterns of metabolic adaptation for preterm and term infants in the first neonatal week. *Archives of Disease in Childhood*, 67(4), 357–365.

Heck, L. J., & Erenberg, A. (1987). Serum glucose levels in term neonates during the first 48 hours of life. *The Journal of Pediatrics*, 110(1), 119–122.

Johnson, T. S. (2003). Hypoglycemia and the Full-Term Newborn: How Well Does Birth Weight for Gestational Age Predict Risk? *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 32(1), 48–57.

Kaiser, J. R., Bai, S., Gibson, N., Holland, G., Lin, T. M., Swearingen, C. J., et al. (2015). Association between transient newborn hypoglycemia and fourth-grade achievement test proficiency: a population-based study. *JAMA Pediatrics*, 169(10), 913–921.

Leuthner, S. R. (2004). Low Apgar scores and the definition of birth asphyxia. *Pediatric Clinics of North America*, 51(3), 737–745.

Low, J. A. (1997). Intrapartum fetal asphyxia: definition, diagnosis, and classification. *American Journal of Obstetrics and Gynecology*, 176(5), 957–959.

Lucas, A., Morley, R., & Cole, T. J. (1988). Adverse neurodevelopmental outcome of moderate neonatal hypoglycaemia. *BMJ*, 297, 1304–1308.

Platt, M. W., & Deshpande, S. (2005). Metabolic adaptation at birth. In *Seminars in Fetal and Neonatal Medicine* 10(4), 341–350.

Rozance, P. J., & Hay Jr, W. W. (2010). Describing hypoglycemia—definition or operational threshold? *Early Human Development*, 86(5), 275–280.

SECTION 4.2

AIDSinfo (2015). Recommendations for use of antiretroviral drugs in pregnant HIV-1-infected women for maternal health and interventions to reduce perinatal HIV transmission in the United States. U.S. <http://www.aidsinfo.nih.gov/contentfiles/lvguidelines/Perinatalgl.pdf>. Accessed 05 Aug 2015.

Barret, B., Tardieu, M., Rustin, P., Lacroix, C., Chabrol, B., Desguerre, I., et al. (2003). Persistent mitochondrial dysfunction in HIV-1-exposed but uninfected infants: clinical screening in a large prospective cohort. *AIDS*, 17(12), 1769–1785.

Barshop, B. A. (2004). Metabolomic approaches to mitochondrial disease: correlation of urine organic acids. *Mitochondrion*, 4, 521–527.

- Blanche, S., Tardieu, M., Benhammou, V., Warszawski, J., & Rustin, P. (2006). Mitochondrial dysfunction following perinatal exposure to nucleoside analogues. *AIDS*, 20(13), 1685–1690.
- Blanche, S., Tardieu, M., Rustin, P., Slama, A., Barret, B., Firtion, G., et al. (1999). Persistent mitochondrial dysfunction and perinatal exposure to antiretroviral nucleoside analogues. *Lancet*, 354, 1084–1089.
- Brogly, S. B., Ylitalo, N., Mofenson, L. M., Oleske, J., Van Dyke, R., Crain, M. J., et al. (2007). In utero nucleoside reverse transcriptase inhibitor exposure and signs of possible mitochondrial dysfunction in HIV-uninfected children. *AIDS*, 21(8), 929–938.
- Bruns, D. E., & Knowler, W. C. (2009). Stabilization of glucose in blood samples: why it matters. *Clinical Chemistry*, 55(5), 850–852.
- Chasela, C. S., Hudgens, M. G., Jamieson, D. J., Kayira, D., Hosseinipour, M. C., Kourtis, A. P., et al. (2010). Maternal or infant antiretroviral drugs to reduce HIV-1 transmission. *The New England Journal of Medicine*, 362(24), 2271–2281.
- Cornblath, M., Hawdon, J. M., Williams, A. F., Aynsley-Green, A., Ward-Platt, M. P., Schwartz, R., et al. (2000). Controversies regarding definition of neonatal hypoglycemia: suggested operational thresholds. *Pediatrics*, 105(5), 1141–1145.
- Cui, X., & Churchill, G. A. (2003). Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology*, 4(210), 1–10.
- De Meyer, T., Sinnaeve, D., Van Gasse, B., Tsiorkova, E., Rietzschel, E. R., De Buyzere, M. L., et al. (2008). NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. *Analytical Chemistry*, 80(10), 3783–3790.
- Doherty, M., Ford, N., Vitoria, M., Weiler, G., & Hirschall, G. (2013). The 2013 WHO guidelines for antiretroviral therapy: evidence-based recommendations to face new epidemic realities. *Current Opinion HIV AIDS*, 8(6), 528–534.
- Dunn, W. B., Bailey, N. J., & Johnson, H. E. (2005). Measuring the metabolome: current analytical technologies. *Analyst*, 130(5), 606–625.

Emwas, A. H., Luchinat, C., Turano, P., Tenori, L., Roy, R., Salek, R. M., et al. (2015). Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review. *Metabolomics*, 11(4), 872–894.

Engelke, U. F. H., Kremer, B., Kluijtmans, L. A. J., van der Graaf, M., Morava, E., Loupatty, F. J., et al. (2006). NMR spectroscopic studies on the late onset form of 3-methylglutaconic aciduria type I and other defects in leucine metabolism. *NMR in Biomedicine*, 19, 271–278.

Engelke, U. F. H., Sass, J. O., Van Coster, R. N., Gerlo, E., Olbrich, H., Krywawych, S., et al. (2008). NMR spectroscopy of aminoacylase 1 deficiency, a novel inborn error of metabolism. *NMR in Biomedicine*, 21, 138–147.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). Multi-and megavariable data analysis part I: basic principles and applications. Umetrics Inc.

Glasgow, A. M., Engel, A. G., Bier, D. M., Perry, L. W., Dickie, M., Todaro, J., et al. (1983). Hypoglycemia, hepatic dysfunction, muscle weakness, cardiomyopathy, free carnitine deficiency and long-chain acylcarnitine excess responsive to medium chain triglyceride diet. *Pediatric Research*, 17, 319–326.

Hawdon, J. M., Ward Platt, M. P., & Aynsley-Green, A. (1992). Patterns of metabolic adaptation for preterm and term infants in the first neonatal week. *Archives of Disease in Childhood*, 67, 357–365.

Ivorra, C., García-Vicent, C., Chaves, F. J., Monleón, D., Morales, J. M., & Lurbe, E. (2012). Metabolomic profiling in blood from umbilical cords of low birth weight newborns. *Journal of Translational Medicine*, 10(1): doi: 10.1186/1479-5876-10-142.

Johnson, M. A., Ahluwalia, G., Connelly, M. C., Cooney, D. A., Broder, S., Johns, D. G., & Fridland, A. (1988). Metabolic pathways for the activation of the antiretroviral agent 2',3'-dideoxyadenosine in human lymphoid cells. *The Journal of Biological Chemistry*, 263(30), 15354–15357.

Kuma, A., Hatano, M., Matsui, M., Yamamoto, A., Nakaya, H., Yoshimori, T., et al. (2004). The role of autophagy during the early neonatal starvation period. *Nature*, 432(7020), 1032–1036.

Levine, B., & Klionsky, D. J. (2004). Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Developmental Cell*, 6(4), 463–477.

McCormack, S. A., & Best, B. M. (2014). Protecting the fetus against HIV infection: a systematic review of placental transfer of antiretrovirals. *Clinical Pharmacokinetics*, 53, 989–1004.

McGarry, J. D., & Foster, D. W. (1980). Regulation of hepatic fatty acid oxidation and ketone body production. *Annual Review of Biochemistry*, 49, 395–420.

McHugh, M. L. (2009). The odds ratio: calculation, usage, and interpretation. *Biochemia Medica*, 19(2), 120–126.

Mikesh, L. M., & Bruns, D. E. (2008). Stabilization of glucose in blood specimens: mechanism of delay in fluoride inhibition of glycolysis. *Clinical Chemistry*, 54(5), 930–932.

Newell, M. L., Brahmbhatt, H., & Ghys, P. D. (2004). Child mortality and HIV infection in Africa: a review. *AIDS*, 18(Supplementary 2), S27–S34.

Newell, M. L., & Bunders, M. J. (2013). Safety of antiretroviral drugs in pregnancy and breastfeeding for mother and child. *Current Opinion HIV AIDS*, 8(5), 504–510.

Newell, M. L. (2006). Current issues in the prevention of mother-to-child transmission of HIV-1 infection. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 100(1), 1–5.

Newell, M. L. (1998). Mechanisms and timing of mother-to-child transmission of HIV-1. *AIDS*, 12(8), 831–837.

Novak, M., Monkus, E. F., Chung, D., & Buch, M. (1981). Carnitine in the perinatal metabolism of lipids I. relationship between maternal and fetal plasma levels of carnitine and acylcarnitines. *Pediatrics*, 67(1), 95–100.

Osborne, J. W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analyses. *Practical Assessment Research & Evaluation*, 11(7), 1–6.

Reinecke, C. J., Koekemoer, G., van der Westhuizen, F. H., Louw, R., Lindeque, J. Z., Mienie, L. J., et al. (2012). Metabolomics of urinary organic acids in respiratory chain deficiencies in children. *Metabolomics*, 8(2), 264–283.

Reinke, S. N., Walsh, B. H., Boylan, G. B., Sykes, B. D., Kenny, L. C., Murray, D. M., et al. (2013). ¹H-NMR Derived Metabolomic Profile of neonatal asphyxia in umbilical cord serum: Implications for Hypoxic Ischemic Encephalopathy. *Journal of Proteome Research*, 12(9), 4230–4239.

Saugstad, O. D. (1975). Hypoxanthine as a measurement of hypoxia. *Pediatric Research*, 9(4), 158–161.

Schicho, R., Shaykhutdinov, R., Ngo, J., Nazyrova, A., Schneider, C., Panaccione, R., et al. (2012). Quantitative metabolomic profiling of serum, plasma, and urine by ¹H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. *Journal of Proteome Research*, 11(6), 3344–3357.

Singh, R., & Cuervo, A. M. (2011). Autophagy in the cellular energetic balance. *Cell Metabolism*, 13(5), 495–504.

Sprecher, S., Soumerknoff, G., Puissant, F., & Degueldre, M. (1986). Vertical transmission of HIV in 15-week fetus. *Lancet*, 2(8501), 288–289.

Stanley, C. A., Rozance, P. J., Thornton, P. S., De Leon, D. D., Harris, D., Haymond, M. W., et al. (2015). Re-evaluating “transitional neonatal hypoglycemia”: mechanism and implications for management. *The Journal of Pediatrics*, 166(6), 1520–1525.

Stover, J., Korenromp, E. L., Blakley, M., Komatsu, R., Viisainen, K., Bollinger, L., et al. (2011). Long-term costs and health impact of continued global fund support for antiretroviral therapy. *PLoS One*, 6(6), 1–7.

Tea, I., Le Gall, G., Küster, A., Guignard, N., Alexandre-Gouabau, M. C., et al. (2012). ¹H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*, 7(1), 1–12.

UNAIDS. (2015). Global AIDS response progress reporting http://www.unaids.org/en/resources/documents/2015/GARPR_2015_guidelines. Accessed 09 Aug 2015.

UNAIDS. (2011). Global plan towards the elimination of new HIV infections among children by 2015 and keeping their mothers alive, 2011–2015. http://www.unaids.org/sites/default/files/media_asset/20110609_JC2137_Global-Plan-Elimination-HIV-Children_en_1.pdf. (2011). Accessed December 29, 2012.

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, K. A., & van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142), 1–15.

Vela, J. E., Miller, M. D., Rhodes, G. R., & Ray, A. S. (2008). Effect of nucleoside and nucleotide reverse transcriptase inhibitors of HIV on endogenous nucleotide pools. *Antiviral Therapy*, 13(6), 789–797.

Walsh, B. H., Broadhurst, D. I., Mandal, R., Wishart, D. S., Boylan, G. B., Kenny, L. C., et al. (2012). The Metabolomic profile of umbilical cord blood in neonatal hypoxic ischaemic encephalopathy. *PLoS One*, 7(12), 1–12.

Wevers, R. A., Engelke, U., Wendel, U., de Jong, J. G. N, Gabreëls, F. J. M., & Heerschap, A. (1995). Standardized method for high-resolution ¹H-NMR of cerebrospinal fluid. *Clinical Chemistry*, 41(5), 744–751.

Wevers, R. A., Engelke, U., & Heerschap A. (1994). High-resolution ¹H-NMR spectroscopy of blood plasma for metabolic studies. *Clinical Chemistry*. 40(7), 1245–1250.

Yu, Z., Kastenmüller, G., He, Y., Belcredi, P., Möller, G., Prehn, C., et al. (2011). Differences between human plasma and serum metabolite profiles. *PLoS One*, 6(7), 1–6.

CHAPTER 5 REPEATABILITY IN THE GENERATION OF A MASS-SPECTROMETRIC-BASED METABOLOMICS DATA MATRIX

5.1 Extended metabolite identification in HIV-ART cord blood of neonates

We explained in Chapter 4 that the statistical data on two metabolites were clear indicators of allostasis associated with transitional neonatal hypoglycaemia (Table 4-2 and Figures 4–3d and 4–3e). The two metabolites were (1) decreased 3-hydroxybutyric acid ($p = 0.007$; ES = 0.49; FC = –2.19) and (2) increased hypoxanthine ($p = 0.023$; ES = 0.40; FC = 3.02). Two additional observations called for further investigations: (1) carnitine from the cord blood of the neonates (exposed and unexposed to HIV-ART) was significantly reduced ($p < 0.0001$) relative to the infant controls. Between the exposed and unexposed neonates the carnitine levels were not significantly different ($p=0.311$) and very low for both groups (Chapter 4: Figure 4-3a). (2) Four amino acids (phenylalanine, glutamine, tyrosine and proline) were included in the list of 11 metabolites associated with spectral bins presenting the most significant measures (VIP ≥ 1.0) for discriminating between infant and unexposed experimental groups (Chapter 4: Table A1-4), although the quantified values were not significantly different ($p > 0.05$) between the exposed and unexposed neonates (Chapter 4: Table 4-2). To investigate these observations further, I performed MS-based metabolomics experiments to elucidate the role of carnitines and amino acids at birth, and on possible effects due to HIV-ART exposure.

In a metabolomics study the method used should be validated in order to regard it as “fit-for-purpose”. Consider the following definitions of repeatability and reproducibility – in metabolomics they are often used when establishing an experimental design. *Repeatability* and *reproducibility* are ways of measuring *precision*. These three terms are described in a review by Naz, Vallejo, Garcia and Barbas (2014). The authors (in their Table 1, which lists parameters for the validation of an analytical method) provide the following definitions:

- “*Precision: The precision of an analytical procedure is defined as the closeness of agreement between a series of analytical measurements obtained from multiple sampling of the same homogenous sample under*

the same developed conditions. Precision can be considered at two levels: repeatability and intermediate precision.” (Naz et al. 2014).

- *“Repeatability: can be defined as the precision under the same operating conditions over a short interval of time, may be performed in the same day. Repeatability is also termed intra-assay precision.” (Naz et al. 2014).* In order for repeatability to be achieved in an experiment, then “conditions must be in the same place, the same location, the same analyst/observer, the same measuring procedures, the same measuring instrument used under the same conditions and repeated over a short period of time” (Van Batenburg et al. 2011)
- *“Reproducibility (Intermediate precision): the precision performed within-laboratories variations: different days, different analysts, and different equipment. This is also termed inter-assay precision.” (Naz et al. 2014)*

High-resolution NMR spectroscopy discussed in Chapter 4 is a quantitative and nondestructive, robust and reliable analytical technique that generates highly reproducible and repeatable metabolomics data (Dumas et al. 2006). Already in the pioneering stages of metabolomics research it was cautioned that metabolomics data produced from gas chromatographic– and liquid chromatographic–mass spectrometric (GC-MS and LC-MS) analysis are not fully reproducible. These methods tend to results with great diversity in chemical composition and wide concentration ranges, which pose a significant challenge as the methods need to be repeatable and reproducible to enable samples to provide reliable information (Glassbrook et al. 2000).

The aims of this chapter are thus:

- (1) To provide a brief oversight of factors contributing to the consistency or lack thereof of semi-targeted metabolomics data.

(2) To indicate a new method, in which I briefly participated, for assessing the analyst's ability to generate a repeatable and reproducible metabolomics data matrix, published as Mason *et al.* 2014.

(3) To provide a detailed description of the analysis of carnitines and amino acids from cord blood using MS-based techniques, with a focus on repeatability in the generation of a metabolomics data matrix.

5.2 Metabolomics data reliability

Metabolomics-driven studies strive to deliver both unbiased qualitative and quantitative metabolite information from acquired biological samples. Therefore, the choice of method for sample preparation is fundamental, since it can affect the overall observed metabolome, and its biological interpretation. In a metabolomics investigation we encounter many different sources of variance in the results such as differences in sample handling, data pre-treatment and pre-processing and, most importantly, instrument reliability (Dunn *et al.* 2005; Gika *et al.* 2007; Crews *et al.* 2009). These sources of variability in a metabolomics investigation need to be carefully assessed and kept to a minimum in order to elucidate the metabolomics information in the data. That is, the variation of interest i.e. variation between groups or individual observations, depending on the experimental design. This is why it is imperative that other sources of variation, biological or analytical, be anticipated and controlled through contingency plans that are implemented in a well-designed experimental strategy.

Metabolites constitute great chemical diversity such as amino and fatty acids, carbohydrates and lipids, so that the analytical equipment currently available is not technologically advanced enough to measure all metabolites with a single universal technique in a single analysis (Naz *et al.* 2014). Since each analytical technique has its limitations, a combination of techniques is required to capture a comprehensive metabolome perspective. Metabolomic profiling regardless of the analytical platform – MS or NMR – requires primarily the generation of a reliable metabolomics matrix for data processing, disclosure of information and eventually its interpretation. This requires a skilled analyst and trustworthy bioinformatic systems that can generate reliable, reproducible information from a metabolomics data matrix.

5.2.1 Generation of a metabolomics matrix

In a metabolomics investigation the raw data have to be exported from the analytical instrument, preprocessed and organized into a table that fully represents the entire analysis and provides for the subsequent statistical analyses. This data table is known as a data matrix as shown in Figure 5-1. In chromatographic data analysis, especially when dealing with LC-MS, the common issue is alignment of the data. A data matrix is the collection of detected and identified variables/metabolites present in a sample, usually normalized against internal standard(s). A data matrix needs to represent each variable across all observations .i.e., the variables need to be consistent (the same) across the entire analysis (Katajamaa & Orešič 2007). Therefore, if metabolites such as alanine, succinic and hippuric acids are identified in a biological sample, their spectral representation must stay the same for other samples analyzed in the same (retention) time slot. Consistency is important in a metabolomics data matrix, otherwise the whole notion of comparing samples based on their chemical changes in metabolite patterns becomes a futile exercise (Jonsson *et al.* 2005). In Figure 5-1 on the far right a representation of a data matrix is presented inside a text-box where n denotes the number of cases and p the number of features/metabolites.

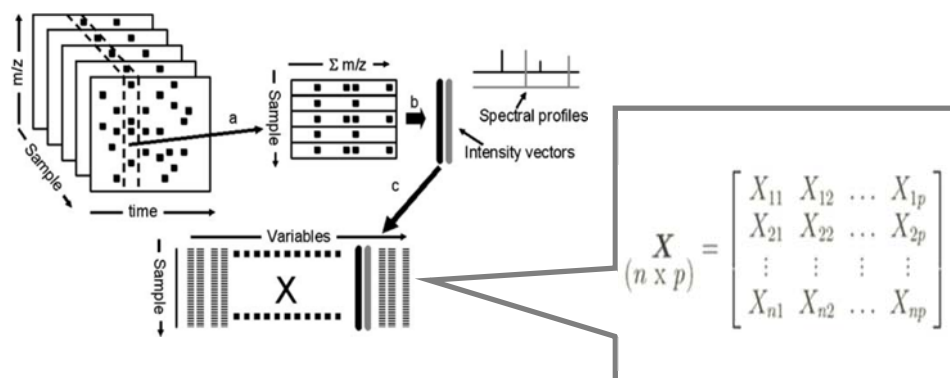


Figure 5–1: Data processing and the creation of a metabolomics data matrix (adapted from Jonsson *et al.* 2005).

5.2.2 The analyst

In metabolomics research quality assurance of the analytical procedures is critical. It is the analyst who plays a central role in ensuring quality by the decision he or she makes (Milman 2005). Advancements in instrumental technology and metabolomic machinery has led to more complex data sets, which results in the need for better processing and analytical software. As a metabolomics researcher sets out to investigate a research question, it is his or her experience and knowledge that helps him or her compile the experimental and measurement design. There are key elements that the analyst considers when setting out on a metabolomics investigation, and these are addressed below.

5.2.2.1 Sample selection and preparation

The first step in generating metabolomics data is sample collection and preparation, a critical step which is often underestimated in metabolomics investigations. In a metabolomics study we need to ensure that the identified metabolites of the collected and extracted sample, represent the metabolome of the same sample prior to collection (Dunn *et al.* 2011).

In a semi-targeted investigation such as presented in this chapter, the choice of sample preparation is dependent on the analytical method and the metabolites of interest. Dajana Vuckovic in his review (2012) mentions that “*an ideal sample-preparation method for global metabolomics should (i) be as non-selective as possible, (ii) be simple and fast to prevent metabolite loss or degradation, (iii) be reproducible and (iv) incorporate a metabolism-quenching step*”.

Two approaches were followed in sample selection as described in this Chapter:

- (1) For the development of the new method to assess repeatability, a *single* sample from a specific biofluid was sub-divided into five identical samples, used for the generation of the data matrix.
- (2) For the investigation of a potential effect of ART on the carnitines, and amino acids, all samples were cord blood from neonates exposed and unexposed (controls) to HIV-ART.

5.2.2.2 Analytical instrumentation: Mass spectrometry

In any chromatographic–mass spectrometry system, the contact between sample and instrument is unavoidable and inevitable. This can cause carry-over contamination of the instrument into the sample, leading to variations in the metabolome irrelevant to the biological system under study. This problem is not encountered in NMR analysis, since there is no contact between instrument and sample. Because not all samples can be run in a single analysis, chromatographic metabolomic studies need to be designed for their experimental purpose, and should include randomization (for both small- and large-scale studies). Randomization must be incorporated in both sample preparation and the order in which samples are injected into the analytical machine. This is to ensure that in the experiment itself there is no bias introduced into the subsequent statistical analysis (Dunn *et al.* 2011). Another limitation encountered in chromatographic experiments is instrument reliability. When using a mass spectrometry technique, the metabolites are separated through chromatographic steps, ionized and detected in the MS. However, the degree and timing of signal attenuation is not the same for all measured analytes, and the biofluid which is measured also plays a role. The likelihood of drift in the chromatographic and MS performance leads to poor repeatability and reproducibility; the analyst must therefore be able to pick up whether there was drift in the analytical run.

5.2.2.3 Quality control

In a GC- or LC-MS system which has relatively long run times, derivatized samples are known to reside in the auto-sampler for quite some time before they are injected into the column, especially so for larger sample cohorts. It is recommended that quality control samples (QCs) be assessed intermittently throughout an analytical run in order to enable robust quality assurance processes to be implemented for each metabolic feature that is detected (Dunn *et al.* 2012). Quality control samples are, in the context of metabolomics studies, “*a representative of the qualitative and quantitative composition of the subject samples being analyzed in the study*” (Dunn *et al.* 2012). These are a pool of samples created from the biofluid test sample obtained for study, in small-scale studies ($n < 100$) such as the one presented below in section 5.5 and in Chapter 6. The preparation of a single pooled QC sample is relatively simple. In our investigation small aliquots of human serum samples were combined together and thoroughly mixed to prepare a pooled QC sample (Dunn *et al.* 2011; Sangster *et al.* 2006).

Dunn *et al.* (2012) described how QC samples should be included in an analytical metabolomics experiment. They advise that “*a typical run order for GC-MS and UPLC-MS, the first n -th injections ($n = 5$ for GC-MS and $n = 10$ for UPLC-MS) are applied to ‘equilibrate’ the analytical system following routine maintenance*”. This is in agreement with Naz *et al.* (2014), who strongly recommend including a number of QCs to be injected right at the beginning of an experiment to monitor the stability of the system, to correct for any machine drift and also to implement QCs right at the end as well, which should be repeatedly regulated at intervals throughout the sequence run (Naz *et al.* 2014). In our laboratory practice we agree with these views, but use a relatively few repeat QC samples. Large numbers of QC samples are not ideal: a tradeoff between time and equilibration, on the one hand, and expense and column contamination on the other, has to be taken into account. Our practical approach is to use only three initial QCs in a batch for GC analysis: the first two are for equilibration and are thus discarded; the third QC is then used to assess the severity of unwanted analytical variation. For our LC-QQQ we used 3–5 QCs in the beginning, depending on the column and the flow rate we use (with UPLC where the flow is high, the column is arguably equilibrated much quicker). This laboratory practice was implemented in my own semi-targeted investigation of cord blood of HIV-ART exposed against ART unexposed neonates (further details regarding this experiment are described in Chapter 6.2).

This chapter continues by illustrating a new approach for assessing the repeatability of the generation of a metabolomics data matrix required for a bioinformatic analysis required for metabolite profiling. Based on mass spectrometry and chromatography techniques, two central experimental questions are addressed in the new method, which are also generally addressed in metabolomics research:

- (1) Assessment of the key aspect in all metabolomics investigations: the requirement that the analyst should produce data of a high quality.
- (2) To develop a method that is an easily applied tool for the convincing holistic qualitative assessment of the repeatability as well as the reproducibility of a metabolomics matrix.

5.3 KEMREP: A new method for the assessment of an analyst's ability

5.3.1 Background

A statistical method to address the two objectives described above was developed by Dr Gerhard Koekemoer, statistician at North-West University. He used the acronym KEMREP (standing for kernel method for the assessment of repeatability and reproducibility) to describe a new method that he developed to address the two problem statements mentioned above. One comment from a reviewer of the first version of the manuscript submitted for publication was:

“This is a paper evaluating a mathematical algorithm for smoothing the GC-MS metabolic data. In the paper, they looked at organic acids in the complex sample, the samples were repeatedly analyzed by two GC-MS and the results/GC spectrum were smoothed/treated using their algorithm based on a mathematical estimation method. Their idea is that by applying this smoothing method they can remove the irreproducible features and generate more reliable profiles. They claim that the method could be used for any analysis but they chose to compare only two very similar GC-MS instruments, both located within their department.”

We introduced various approaches to address the above concern: (1) Three analysts participated in repeat analysis for the revised manuscript. (2) A strict standard operational procedure was developed to generate the repeat metabolomics data. (3) We used different biofluids for the repeat analysis. (4) We applied GC- and LC-based metabolomics. My contribution was to generate GC-MS metabolomics data from a urine sample, also used by two other analysts, for assessment using KEMREP.

The revised and published manuscript on KEMREP (see Annexure 2) using GC-MS indicates quantitatively whether the experiment conducted by a researcher is repeatable and reproducible. Figure 5-2 is a workflow diagram, which indicates the experimental procedures that we followed to generate data for KEMREP: (1) The first experiment was based on GC-MS (section 5.3), which used urine as a biofluid. This experiment aimed at determining the *inter*-analyst reproducibility against *intra*-analyst repeatability. [It also measured reproducibility as it involved 3 different analysts who performed the same experiment on the same sample of 5 repeats.] (2) The second experiment used an LC-MS system (section 5.4). This procedure aimed to determine qualitatively the repeatability of the analysts' laboratory and analytical skills.

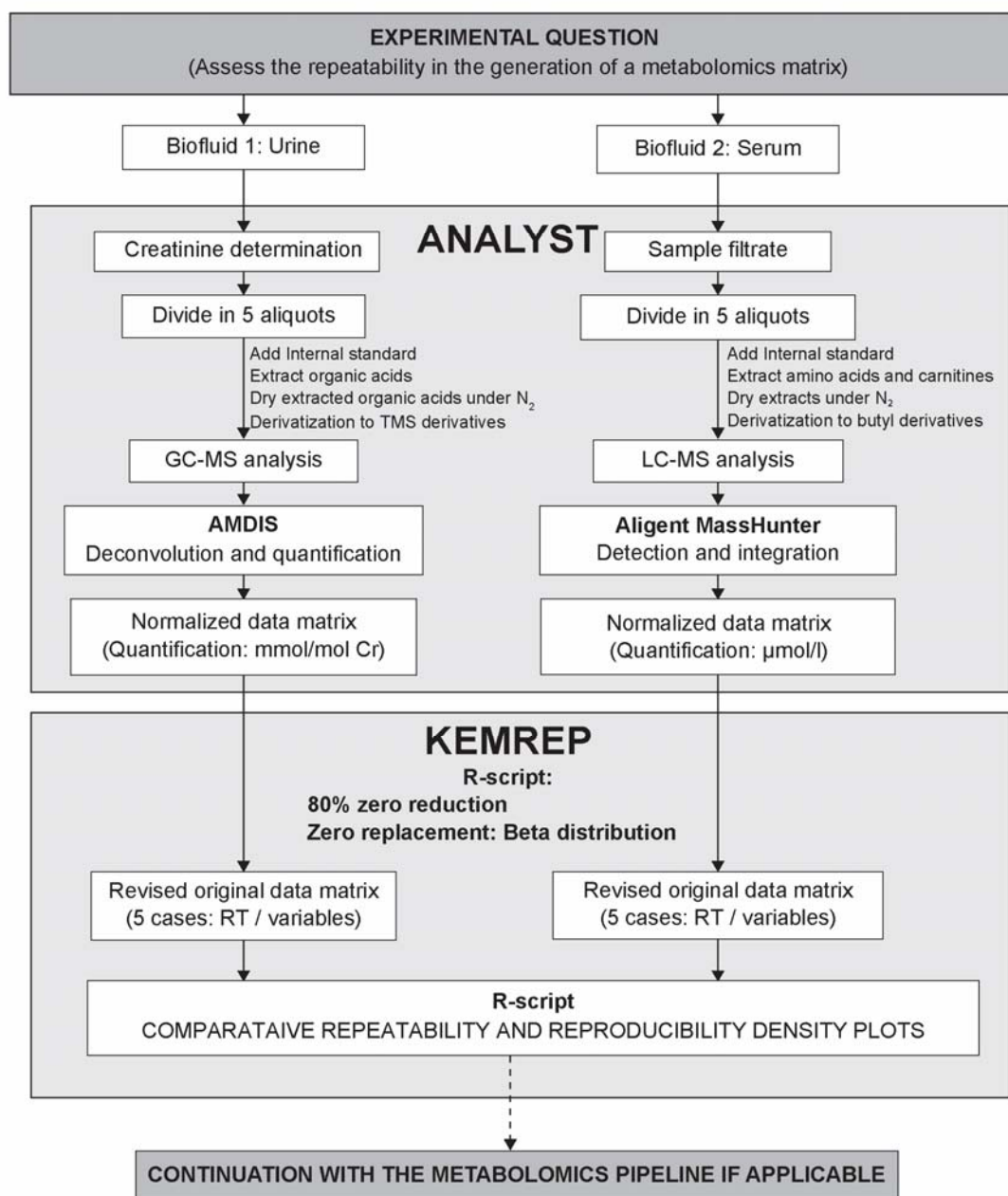


Figure 5–2: Flow diagram illustrating the experimental procedures followed by analysts in the metabolomics pipeline and showing the role subsequently played by KEMREP.

Figure 5-2 indicates the procedures followed to generate the data. Samples from two biofluids [urine from a healthy adult and cerebrospinal fluid (CSF) from a patient infected with, and treated for, tuberculosis meningitis (TBM)] were selected for assessment of KEMREP. We used aliquots from the same sample for analysis by the

same or different analysts, using the same experimental protocol although generating the data on different GC-MS platforms, to address the issue of reproducibility.

5.3.2 Experimental method and results

A urine sample was obtained from a single healthy male participant. A sample of approximately 20 mL was collected and 0.6 mL was used for creatinine determination. Five aliquots of 3 mL were measured out for further repeatability assessment based on organic acid analysis, in the generation of a metabolomics data set. Further experimental details are described in Annexure 2, and the results of my GC-MS contribution, taken from the published KEMREP paper, is indicated in Figure 5-3.

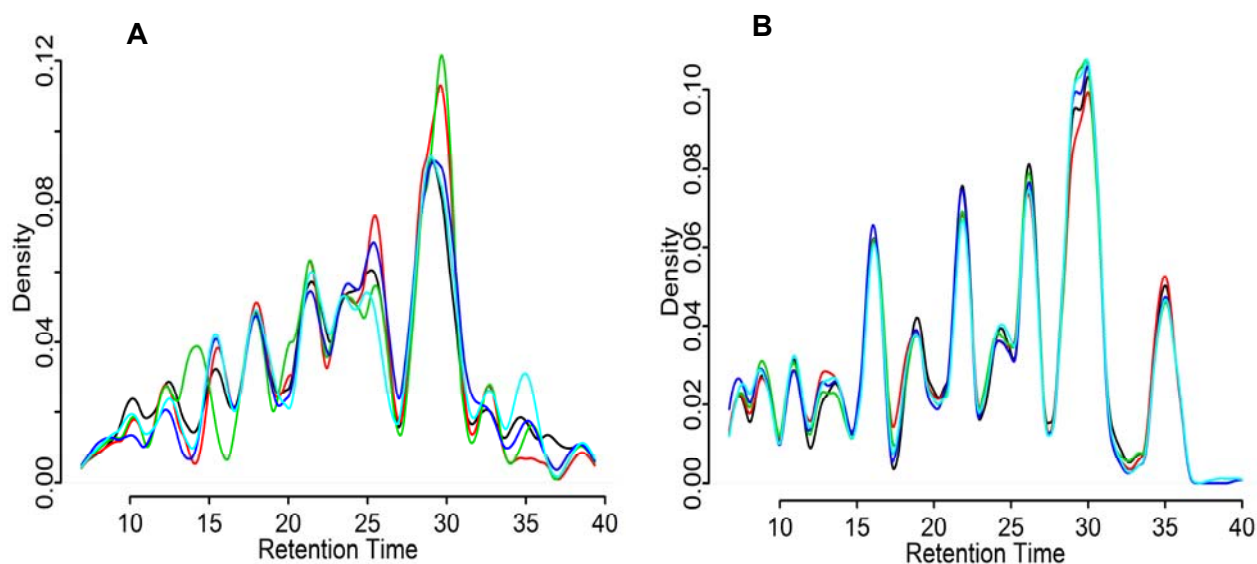


Figure 5–3: Overlays of density plots of 5 repeat aliquots from a single urine sample. (A) First attempt to generate the GS-MS data. (B) Second attempt to generate the GS-MS data by the same analyst. All analyses were conducted using the identical reagents, standard operating procedure (SOP) and the same GC-MS apparatus.

The repeatability/reproducibility of a metabolomics matrix generated by application of the KEMREP method in these cases is shown in Figure 6 of the KEMREP publication (Annexure 2). In Figures 6A, 6B and 6C (Annexure 2) the smoothed chromatographic profiles after the first attempt by the three different analysts to produce a metabolomics data matrix using the same SOP is shown, as well as Figures 6D, 6E and 6F (Annexure 2) that represent the corresponding results of a second round of analysis by each analyst. My own contribution as one of the three analysts is shown in Figure 5-3 (A and B). A visual, qualitative assessment of the result from my first analysis is shown in Figure 5-3A, which clearly indicates insufficient repeatability. (This was the case for each of the analysts, which indicated insufficient reproducibility overall. This matter is not further elaborated on here.)

For the second attempt, some key features of organic acid analysis had to be adhered to meticulously by the three analysts. These include (1) the use of freshly prepared reagents; (2) glassware and Hamilton syringes cleaned in exactly the same manner; (3) the analyses done in the same laboratory and using the same GC-MS equipment, and finally (4) generating and analyzing the raw data using the same AMDIS settings. The qualitative repeatability and reproducibility profiles in all cases improved substantially at the second attempt, and the outcome of my second attempt (Figure 5-3B) produced high repeatability. These results clearly substantiate the applicability of the KEMREP method to highlight the central role of the analysts in research or routine experiments and for a way to effectively communicate the qualitative status of metabolomics data generated by such experts.

5.3.3 The KEMREP publication

The revised manuscript was accepted for publication in *Current Metabolomics* (Vol. 2, pp. 15-26 (2014) plus supplementary information), shown in SI, Annexure 2. The title page and Abstract are shown below.

KEMREP: A New Qualitative Method for the Assessment of an Analyst's Ability to Generate a Metabolomics Data Matrix by Gas Chromatography–Mass Spectrometry

Shayne Mason¹, Gontse P. Moutloatse¹, A. Marceline van Furth², Regan Solomons³, Mari van Reenen⁴, Carolus Reinecke^{1*} and Gerhard Koekemoer⁴

¹Centre for Human Metabonomics, Faculty of Natural Sciences, Private Bag X2006, North-West University (Potchefstroom Campus), South Africa; ²Department of Paediatric Infectious Diseases–Immunology and Rheumatology, Vrije Universiteit Medical Centre, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands; ³Department of Paediatrics and Child Health, Tygerberg Hospital, University of Stellenbosch, PO Box 19063, Tygerberg 7505, South Africa; ⁴Statistical Consultation Services, Faculty of Natural Sciences, Private Bag X2006, North-West University (Potchefstroom Campus), South Africa

Abstract: The analytical procedures required to generate a quantified metabolomics data matrix include many and widely different potential sources of error, complicating the generation of reliable data. The methods generally used to assess precision of such data all have distinct merits but some clear limitations as well. In this paper we describe KEMREP (kernel method for the assessment of repeatability and reproducibility), a new method with the advantage and focus aimed specifically at analysis of the reliability of metabolomics data. Repeatability and reproducibility were assessed on gas chromatography-mass spectrometry (GC-MS) generated metabolomics data matrices produced by and between analysts and across laboratories, using cerebrospinal fluid (CSF) and urine as biological samples for analysis. KEMREP provides a visual overlay of the smoothed and scaled versions of the data from repeated samples for a direct and easy qualitative assessment of repeatability or reproducibility of a distinct chromatographic region (univariate) or for the experiment as a whole (multivariate). The KEMREP method can also be extended by the imposition of confidence bounds which provide lower and upper limits that indicate quantitatively whether the experiment was repeatable or reproducible at a predefined input coefficient of variation (CV). KEMREP is thus a novel approach which supplements existing methods of assessment of reliability of metabolomics data; provides a benchmark for assessing the quality of practical work performed by analysts; monitors the sequence of data pre-treatment steps; and tests the robustness of an experimentally designed protocol for metabolomics.

Keywords: Gas chromatography-mass spectrometry, kernel density, metabolomics, qualitative, repeatability, reproducibility.

5.4 Application of KEMREP in LC-QQQ

5.4.1 Experimental approach

When using LC-MS, non-targeted analysis is the most commonly used method. In the following repeatability study, two serum samples were collected but analyzed on different days: (1) a healthy individual – control labelled sample Day 1 [AM] (see Table 5-1), and (2) a patient who presented with phenylketonuria (PKU) labelled Day 3 [P] (see Table 5-1). The samples were aliquoted into five equal-volume repeats (10 μ L) and analyzed with an LC-QQQ system. The primary aim was to assess the

repeatability of an amino acid and an acylcarnitine analysis using electrospray ionization tandem mass spectrometry (ESI-MS/MS, abbreviated as high-performance triple quadrupole mass spectrometry, i.e. LC-QQQ in the thesis). The outline of the method used for this repeatability study is the same as used for the development of KEMREP, as presented in Figure 5-2.

5.4.2 Reagents

Reagents and standards for the analysis of carnitines: acetonitrile, formic acid and methanol were purchased from Merck Chemical Co. (Darmstadt, Germany). Acetic acid, valine, leucine, isoleucine, phenylalanine, methionine, citrulline, glycine and lysine were purchases from Sigma-Aldrich Co. The following carnitine and acylcarnitine standards and deuterated carnitine and acylcarnitine standards were obtained from Dr H. J. ten Brink, Free University Hospital (Amsterdam, The Netherlands): L-carnitine.HCl, acetyl-L-carnitine.HCl, propionyl-L-carnitine.HCl, isovaleryl-L-carnitine.HCl, octanoyl-L-carnitine.HCl, hexadecanoyl-L-carnitine.HCl, [methyl-d3]-L-carnitine.HCl, [d3]-acetyl-L-carnitine.HCl, [3,3,3-d3]-propionyl-L-carnitine.HCl, [d9]-isovaleryl-L-carnitine.HCl, [8,8,8-d3]-octanoyl-L-carnitine.HCl and [16,16,16-d3]-hexadecanoyl-L-carnitine.HCl. The following deuterated amino acids were obtained from Cambridge Isotope Laboratories Inc. (Andover, MA, USA): [d10]-L-isoleucine, [d8]-L-valine, [d2]-glycine, [d3]-methyl-L-methionine, [d5]-ring-L-phenylalanine, [d5]-L-glutamine, [d5]-indole-L-tryptophan, [d4]-L-lysine.2HCl and [d4]-L-citrulline. [For concentrations of deuterated carnitines see Annexure 3]

5.4.3 Sample preparation

The LC-QQQ method for the profiling of serum acylcarnitines and amino acids as described in Rashed *et al.* 1995; Mels *et al.* 2011) was used with minor modifications to serum samples in this repeatability investigation. Five sample repeats of 10 µL serum were aliquoted into 1.5 mL microcentrifuge tubes followed by the addition of 410 µL ice-cold isotope (labelled internal standard [deuterated acylcarnitines and amino acids]) mixture prepared in methanol. The concentrations of the stable isotopes and their preparation are presented in Table A3-3 in Annexure 3. The samples were vortexed briefly and centrifuged for 10 minutes at 6 000 rpm. The supernatant was carefully removed without disturbing the protein pellet and transferred to a 2 mL GC-MS vial. The samples were dried under a gentle stream of nitrogen (at 65 °C) for 15 minutes.

The residue of the dried samples was derivatised with the addition of 100 µL butanolic-HCl (3N) (prepared fresh). Samples were capped, briefly vortexed and incubated for 15 minutes at 65 °C, followed by drying under nitrogen at 65 °C. Butylated residues were re-dissolved in 100 µL acetonitrile:water (50:50 v/v in 0.1% formic acid) mixture and transferred to vial inserts before LC-QQQ analysis.

For the repeatability analysis, two types of blank samples were introduced: (1) A blank sample which contained no sample or isotope solution, but underwent butylation. This blank sample was always loaded first in the run order of the LC-QQQ work-list. (2) Additionally, two blank samples which contained only the labelled internal standard isotopes also underwent the entire same derivatization procedure as the sample material. These isotope-blanks were loaded in the second and third positions of the analytical run and, if QC samples were created, they were placed after the isotope-blanks.

5.4.4 LC-QQQ acylcarnitine and amino acid analysis

The LC-QQQ method used for the separation and identification of acylcarnitine as described by Ferrer *et al.* (2007) and Mels *et al.* (2011) was used with slight adjustments to separate butyrylcarnitine and isobutyrylcarnitine. LC-QQQ analysis was performed on a 1200 series Agilent HPLC system linked to an Agilent 6410 MS/MS system. A C18 Zorbax SB-Aq column (150 mm × 2.1 mm × 3.5 µm) from Agilent was used for chromatographic separation of butylated compounds and kept at 30 °C throughout the analysis. Samples in the auto-sampler were kept at 4 °C. One µL sample was injected followed by the initiation of the programmed solvent gradient. The gradient for mobile phases A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile) is shown in Table A3-6, Annexure 3. Each run lasted 23 minutes followed by a 7-minute post-run to re-equilibrate the column. The gas temperature for the ESI was kept at 300 °C at a flow rate of 7.5 L/minute. The capillary voltage was 3500 V and the nebulizer pressure was set at 30 psi. The MRM transitions used to detect the target amino acids and the outcome of the analysis is shown in

5.4.5 Data extraction and normalisation of acylcarnitines and amino acids

Agilent's MassHunter Quantitative Analysis software was used for detection and integration of the compound peaks according to the manufacturer's specifications. The integration of all detected peaks was checked for correctness. The integrated-peak data table was then exported to MS Excel where normalization (relative quantification)

of the compounds was performed. The compounds which had related stable isotope internal standards were normalized with their respective isotopes. For those compounds that did not have related isotopes, isotopes that were structurally very similar and eluted closest were used to normalize these compounds.

5.4.6 Statistical analysis

Henceforth we will discuss the figure of merit and statistical method used to assess analytical repeatability (i.e. assuming the same analyst, equipment, and laboratory). We chose to use the coefficient of variation (CV) as a figure of merit, since it is traditionally applied in this context, and the KEMREP paper (Mason *et al.* 2014).

Presented below in Figure 5-4 are the repeatability results of the semi-targeted LC-QQQ experiment to test the analyst's experimental repeatability. CV percentages are calculated as a ratio of the standard deviation divided by the mean multiplied by 100 and plotted against the mean values for each feature (metabolite). The red lines in the CV graphs (Figures 5-4a and 5-4c) represent an acceptable threshold for this figure of merit. CV values below 30% are deemed acceptable. This rule of thumb was suggested by the US Food and Drug Administration (FDA), since there are no generally accepted standards for the assessment of repeatability in metabolomic data sets. Note that CV values are calculated as 100 times the RSD value, to express them as a percentage, hence these measures are essentially the same. The RSD affords a univariate measure of intra-experiment precision for every variable/metabolic feature separately and it regulates data for comparison with acceptance criteria.

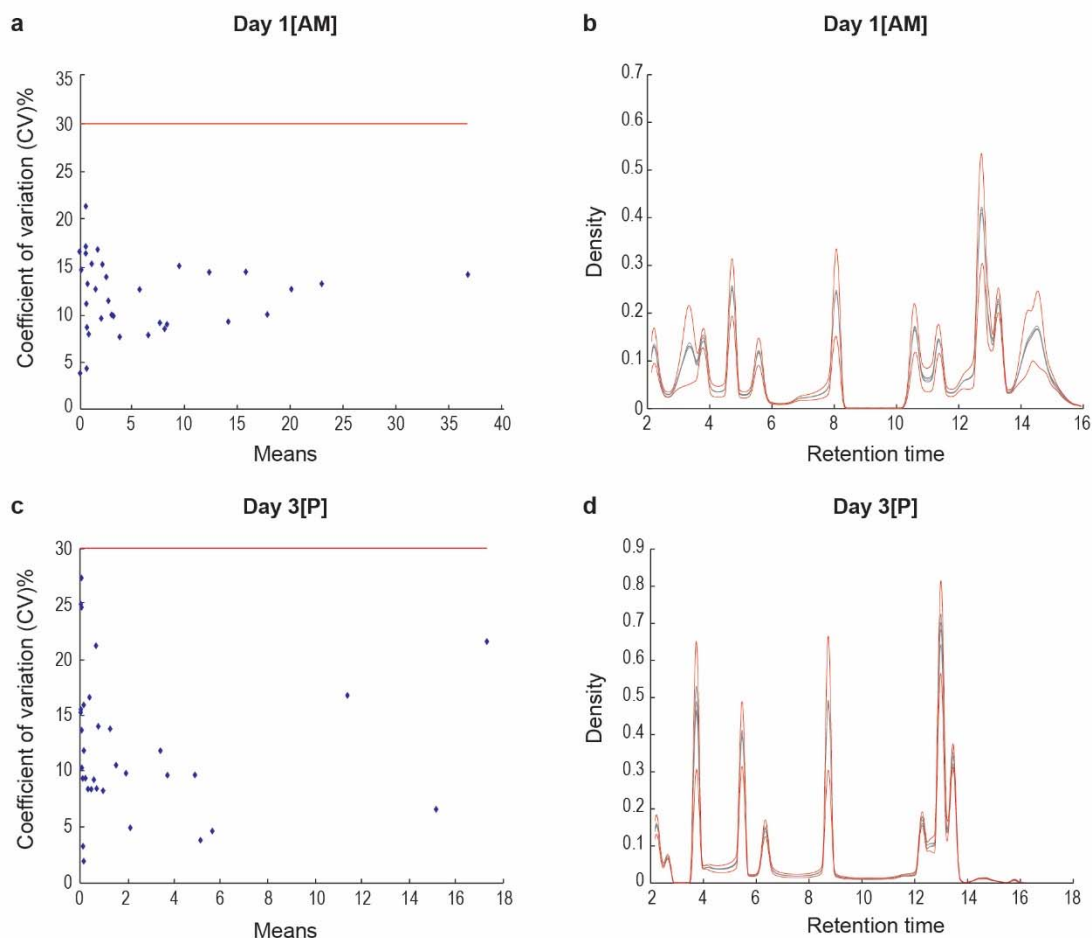


Figure 5-4: Repeatability results: CV distribution, density methods and kernel density estimates of an LC-QQQ acylcarnitine and amino acid analysis done on two days. Day 1 repeatability analysis is represented in 5-4a and 5-4b. Following Day 3 analysis are presented in 5-4c and 5-4d.

Dunn *et al.* (2012) outlined the FDA's criteria range as

“the precision of an analytical method describes the closeness of individual measures of an analyte when the procedure is applied repeatedly to multiple aliquots of a single homogenous volume of biological matrix. Precision should be measured using a minimum of five determinations per concentration. The precision determined at each concentration level should not exceed 15% of the coefficient of variation (CV) except for the low limit of quantification (LLOQ), where it should not exceed 20% of the CV.”

For the application of non-targeted metabolomic analysis, such as LC-QQQ, CV values close to 30% are acceptable (Naz *et al.* 2014; Guy *et al.* 2008). In our investigation we applied LC-QQQ for metabolic profiling which was not specific to one analyte but a comprehensive analysis of the entire metabolome, thereby detecting hundreds of analytes and so an acceptance tolerance of 30% is allowed. The resulting CVs reported in Figure 5-5a Day 1 (labeled as [AM]) show a maximum CV of 22% and 25% for Day 3 (labeled as [P]) samples (Figure 5-4c), both acceptable (i.e. below 30%) and therefore indicating good repeatability.

In Figure 5-4, the data were analyzed using the kernel density method as described in the KEMPREP manuscript (see Annexure 3). From the smoothed chromatogram we were able to assess qualitatively if the experiment was repeatable. Overlapping estimates are indicative of minimal variation, which implies that the experiment was repeatable. Since slight variation can be expected, a confidence interval (CI) was constructed to assess whether the variability was within acceptable bounds. These bounds (or CIs) are represented by the red lines in Figure 5-4b and d. Results of both Day 1 and Day 3 smoothed chromatograms of Figure 5-4b and d, indicate repeatable results with the smoothed chromatograms of both data sets overlapping. From these results it is clear that both analytical runs were repeatable given that the smoothed chromatograms fell within the CI – their analysis was therefore successfully repeatable.

5.4.7 Conclusions

From the results presented in this Chapter, the following conclusions can be reached on the experimental aims defined in section 5.1:

The achievements reported in this chapter are thus:

(1) A new method for assessing the analyst's ability to generate a repeatable and reproducible metabolomics data matrix was successfully developed (Mason *et al.* 2014). My contribution provided an example to illustrate the applicability of the KEMREP method to highlight the central role of the analyst in research or routine experiments, and for a way to effectively communicate the qualitative status of metabolomics data generated by such an analyst.

(2) The results summarized in Figure 5-4 clearly indicate the repeatability of generating a metabolomics matrix for profiling carnitines and amino acids. These results substantiate the observation that the LC-QQQ/MS/MS system offers high-precision

and repeatable-targeted metabolite profiles with a wide dynamic range and a relatively easy quantification method.

(3) The LC-QQQ/MS/MS approach, evaluated in this Chapter, thus provides the method of choice for the investigation of a possible influence of HIV-ART exposure on the acylcarnitine and amino acid profiles in the cord blood of newborns (Chapter 6).

REFERENCES

- Crews, B., Wikoff, W. R., Patti, G. J., Woo, H. K., Kalisiak, E., Heideker, J., et al. (2009). Variability analysis of human plasma and cerebral spinal fluid reveals statistical significance of changes in mass spectrometry-based metabolomics data. *Analytical Chemistry*, 81(20), 8538–8544.
- Dumas, M. E., Maibaum, E. C., Teague, C., Ueshima, H., Zhou, B., Lindon, J. C., et al. (2006). Assessment of analytical reproducibility of ¹H NMR spectroscopy based metabonomics for large-scale epidemiological research: the INTERMAP Study. *Analytical Chemistry*, 78(7), 2199–2208.
- Dunn, W. B., Bailey, N. J., & Johnson, H. E. (2005). Measuring the metabolome: current analytical technologies. *Analyst*, 130(5), 606–625.
- Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., et al. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*, 6(7), 1060–1083.
- Dunn, W. B., Wilson, I. D., Nicholls, A. W., & Broadhurst, D. (2012). The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis*, 4(18), 2249–2264.
- Ferrer, I., Ruiz-Sala, P., Vicente, Y., Merinero, B., Pérez-Cerdá, C., & Ugarte, M. (2007). Separation and identification of plasma short-chain acylcarnitine isomers by HPLC/MS/MS for the differential diagnosis of fatty acid oxidation defects and organic acidemias. *Journal of Chromatography B*, 860(1), 121–126.
- Gika, H. G., Theodoridis, G. A., Wingate, J. E., & Wilson, I. D. (2007). Within-day reproducibility of an HPLC-MS-based method for metabonomic analysis: application to human urine. *Journal of Proteome Research*, 6(8), 3291–3303.
- Glassbrook, N., Beecher, C., & Ryals, J. (2000). Metabolic profiling on the right path. *Nature Biotechnology*, 18(11), 1142–1143.
- Guy, P. A., Tavazzi, I., Bruce, S. J., Ramadan, Z., & Kochhar, S. (2008). Global metabolic profiling analysis on human urine by UPLC–TOFMS: issues and method validation in nutritional metabolomics. *Journal of Chromatography B*, 871(2), 253–260.

Jonsson, P., Bruce, S. J., Moritz, T., Trygg, J., Sjöström, M., Plumb, R., et al. (2005). Extraction, interpretation and validation of information for comparing samples in metabolic LC/MS data sets. *Analyst*, 130(5), 701–707.

Katajamaa, M., & Orešič, M. (2007). Data processing for mass spectrometry-based metabolomics. *Journal of Chromatography A*, 1158(1), 318–328.

Mason, S., P Moutloatse, G., Marceline van Furth, A., Solomons, R., van Reenen, M., Reinecke, C., & Koekemoer, G. (2014). KEMREP: A new qualitative method for the assessment of an analyst's ability to generate a metabolomics data matrix by gas chromatography–mass spectrometry. *Current Metabolomics*, 2(1), 15–26.

Mels, C., Jansen van Rensburg, P., van der Westhuizen, F. H., Pretorius, P. J., & Erasmus, E. (2011). Increased excretion of C4-carnitine species after a therapeutic acetylsalicylic acid dose: Evidence for an inhibitory effect on short-chain fatty acid metabolism. *ISRN Pharmacology*, 2011, 1–8.

Milman, B. L. (2005). Identification of chemical compounds. *Trends in Analytical Chemistry*, 24(6), 493–508.

Naz, S., Vallejo, M., García, A., & Barbas, C. (2014). Method validation strategies involved in non-targeted metabolomics. *Journal of Chromatography A*, 1353, 99–105.

Rashed, M. S., Ozand, P. T., Bucknall, M. P., & Little, D. (1995). Diagnosis of inborn errors of metabolism from blood spots by acylcarnitines and amino acids profiling using automated electrospray tandem mass spectrometry. *Pediatric Research*, 38(3), 324–331.

Sangster, T., Major, H., Plumb, R., Wilson, A. J., & Wilson, I. D. (2006). A pragmatic and readily implemented quality control strategy for HPLC-MS and GC-MS-based metabonomic analysis. *Analyst*, 131(10), 1075–1078.

Van Batenburg, M. F., Coulier, L., van Eeuwijk, F., Smilde, A. K., & Westerhuis, J. A. (2011). New figures of merit for comprehensive functional genomics data: the metabolomics case. *Analytical Chemistry*, 83(9), 3267–3274.

Vuckovic, D. (2012). Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Analytical and Bioanalytical Chemistry*, 403(6), 1523–1548.

CHAPTER 6 METABOLIC RISKS AT BIRTH FOLLOWING HIV-ART EXPOSURE: INSIGHTS FROM SEMI-TARGETED METABOLOMICS STUDIES

6.1 A semi-targeted metabolomics analysis of cord blood acylcarnitines

6.1.1 Introduction

Metabolic pathways of intermediary metabolism involving processes such as fatty acid oxidation and organic acid metabolism all depend on a functional OXPHOS system (Chapter 2.6.1). NRTI-associated mitochondrial dysfunction is understood to be caused by the inhibition of polymerase gamma (POLG) (Lewis *et al.* 2006). Although, it has been speculated that the length of pregnancy is too short a time for adverse effects of NRTI-induced mitochondrial dysfunction to present themselves at birth; neonates exposed to HIV-ART present with abnormal mitochondrial function, manifested – amongst other conditions – as lactic acidosis (Alimenti *et al.* 2003). In our own MTCT-NMR investigation reported in Chapter 4, we observed high levels of lactic acid ($p = 0.0001$; FC = 3.14), but also decreased carnitine ($p = 0.0001$; FC = -2.46) (Moutloatse *et al.* 2016). The presence of depleted carnitine in our study is described as being due to the reduction in intrauterine maternal-fetal transfer, which has been reported in late gestation (Novak *et al.* 1981; Glasgow *et al.* 1983). However, in disturbances in the OXPHOS system, secondary products of intermediary metabolism, including acylcarnitines, are known to occur (Krivitzky *et al.* 2009; Wilcken 2010). The use of tandem mass spectrometry (MS/MS), based on these observations, has grown to be very popular over the years, especially since its application in the screening of inborn errors of metabolism (IEM). MS/MS offers high-sensitivity, high-precision and repeatable targeted metabolite profiles (Tsugawa *et al.* 2014). It is able accurately to detect products such as acylcarnitines, while indicating abnormal levels of the activities of metabolic pathways such as dysfunctional fatty acid oxidation, organic acids or amino acid metabolism (Wilcken *et al.* 2003), offering a platform for the quantification of such substrates.

An approach to link results from the national–US programme on newborn screening for inborn errors of metabolism, to supposed mitochondrial metabolic disturbances due to HIV-ART, was recently carried out in New York (Kirmse *et al.* (2013). These researchers hypothesized that *“HIV-exposed and ART-exposed children would have disruption of metabolic pathways that are normally dependent on functional OXPHOS, leading to a greater incidence of positive newborn metabolic screens compared with the general newborn population”*. Kirmse *et al.* (2013) reported that the prevalence of abnormal newborn metabolic screens in HIV screen-positive infants was higher than that in the general population (2.2% versus 1.2%, $p = 0.00025$), most of which were for disorders of mitochondria-related metabolism, such as medium-chain acyl-CoA dehydrogenase deficiency (MCADD), multiple acyl-CoA dehydrogenase deficiency (MADD), carnitine palmitoyltransferase deficiency (CPT), maple syrup urine disease (MSUD) and organic acidemias. Abnormal acylcarnitine levels occurred more frequently in ARV-exposed compared to ARV-unexposed infants (43% versus 0%, $p = 0.02$). In contrast to our observation of serum values of carnitine at birth, the mean of free carnitine levels in plasma of the ARV-exposed infants and in the ARV-unexposed infants from the screening analysis (normally blood samples taken 2–3 days after birth) were not significantly different [$10.99 \mu\text{mol/mL}$ vs $12.91 \mu\text{mol/mL}$, $p = 0.89$]. At that stage of life, acylcarnitine levels due to a metabolic disorder were evident:

“In the ARV-exposed group, there were 10 measurements spread over 7 subjects that were >99th percentile for age-matched reference values for 5 different acylcarnitine species (octenolcarnitine (C8:1), palmitoleylcarnitine (C16:1), palmitoylcarnitine (C16), oleoylcarnitine (C18:1) and linoleylcarnitine (C18:2)). Of these, abnormal levels of C8:1 (3 subjects) and C18:1 (3 subjects) represented the most consistent abnormalities detected. All acylcarnitine species that were abnormal corresponded to fatty acid oxidation (even-carbon number) and not organic acid metabolism (short-chain, odd-carbon number)”.

These observations set a line of thinking (Kirmse *et al.* 2013; Jao & Abrams 2014; Jao *et al.* 2016) that acylcarnitine and amino acid products of intermediary metabolism, are increased in ARV-exposed infants (indicated by Jao *et al.* 2016 as 43% vs 0%, $p \ll 0.02$). In terms of monitoring the long-term health of HIV-ART-exposed children, the option of record linkage studies (e.g. as used by Kirmse *et al.* 2013) is seen to deserve consideration, but the feasibility and acceptability of such approaches need to be evaluated within the context of national health policies (Thorne & Tookey 2016).

Against this background, we sought to conduct a follow-up on the observed carnitine results in our NMR investigation on cord blood, by using a liquid chromatography–tandem mass spectrometric (LC-QQQ) system for the analysis of acylcarnitines and analysis of amino acids in the same chromatographic run (see Chapter 5). In view of the highly repeatable results shown in Figure 5-4, it is evident that a semi-targeted LC-QQQ method was our choice of analysis.

6.1.2 Materials and methods

6.1.2.1 Reagents

The reagents and standards used for the determination of acylcarnitine and amino acids in cord blood serum of HIV-ART exposed and unexposed infants is the same as those used and mentioned in Chapter 5, section 5.5.1.1 (a repeatability study which used serum from healthy controls).

6.1.2.2 The experimental subjects for the metabolomics analysis

The participants for this experimental investigation were HIV-ART-exposed, but uninfected neonates (n = 12) against their HIV-ART unexposed controls (n = 15). These cases were from the same cohort used in the MTCT-NMR study except for the PLIEM infant control group that was not included in this investigation. The exposed and unexposed groups' clinical descriptions and sample collection procedures are explained in Chapter 4, section 4.2 (also see Annexure 4). Ethical approval for collecting the samples from the two neonatal groups was obtained from the Ethics Committee of the University of Amsterdam (see Moutlootse *et al.* 2016).

6.1.2.3 Analysis of acylcarnitines using liquid chromatography–tandem mass spectrometry

The procedure used for sample preparation of the cord blood serum for the acylcarnitine analysis, is the ESI-MS/MS method for determination of serum acylcarnitines as described by Mels *et al.* (2011) and Rashed *et al.* (1995). This technique (as described in Chapter 5; section 5.5), was slightly modified to determine acylcarnitines in cord blood serum of HIV-ART-exposed neonates in contrast to their unexposed counterparts.

6.1.2.4 Data processing and statistical analysis

The first procedural step we undertook for our data processing, involved the inspection of blank and QC samples for contamination and chromatographic carry-over. In this investigation however, the method was well standardized and so no carry-overs were observed. The integration of all detected peaks was checked for correctness. The integrated-peak data table was then exported to MS Excel where normalization (relative quantification) of the compounds was performed. The compounds which had related stable isotope internal standards were normalized with their respective isotopes. For those compounds that did not have related isotopes, isotopes that were structurally very similar and eluted closest to them were used to normalize these compounds. Acylcarnitines were quantified by comparison with the signal intensity of carnitine and acylcarnitines against the signal intensity of the corresponding deuterated analogues. The concentrations of analysed carnitine and acylcarnitines were expressed as $\mu\text{mol/L}$.

Statistical analyses were performed using Matlab (MATLAB with Statistics and PLS Toolbox Release (2012b), The MathWorks, Inc., Natick, MA, USA); SAS (SAS Institute Inc. (2015). The SAS System for Windows Release 9.3 TS Level 1M0 Copyright© by SAS Institute Inc., Cary, NC, USA); and SPSS (SPSS Inc. (2015). IBM SPSS Statistics Version 22, Release 22.0.0, ©IBM Corporation and its licensors. <http://www-01.ibm.com/software/analytics/spss/>). Statistical analysis was performed using the same method as the one reported in Chapter 4 – see Moutloatse *et al.* 2016 manuscript, including detection of outliers.

6.1.3 Results and discussion

The unsupervised PCA based on acylcarnitine data from the uninfected infants exposed *in utero* to HIV-ART against their unexposed controls is shown on the left of Figure 6-1. The PCA score plot Figure 6-1a indicated no separation and limited differentiation between the groups, with the first principal component (PC1) accounting for 43% and PC2 for 15% of the variability in the data. The supervised PLS-DA data are presented in Figure 6-1b; it too indicates no separation and a somewhat improved differentiation between the HIV-ART-exposed infants compared to the unexposed cases. Further univariate analysis was applied to the acylcarnitine data (Table 6-1) of which 2 metabolites from the 11 carnitines were revealed to be responsible for the small differentiation between the two experimental groups. Both these acylcarnitines – decanoylcarnitine (C10; $p = 0.030$; $FC = -1.32$) and dodecanoylcarnitine (C12; $p =$

0.037; FC = -1.44) – were decreased in the HIV-ART=exposed group and both are related to medium-chain fatty acids (C8–C12). The mean concentration of the remaining acylcarnitine from this group of fatty acids (octanoylcarnitine, C8) did not significantly differ ($p = 0.068$; FC = -1.28) between the exposed and unexposed groups. In a merged data set of HIV subjects ($n = 32$) and healthy controls ($n = 20$) for pathway mapping and correlation analysis, 58 metabolites distinguished between HIV subjects and controls. Carnitine was decreased in the HIV patients ($p \leq 0.0001$), stearoylcarnitine ($C = 18$; $p = 0.037$) increased and decreased in two short-chain acylcarnitines [butyrylcarnitine ($C = 4$; $p = 0.037$) and propionylcarnitine ($C = 3$; $p = 0.003$)]. In a metabolomics study on healthy subjects, an acylcarnitine pattern dominated by C8:0-, C10:0-, and C12:0-carnitine was identified. The increased acylcarnitines were considered to be prominent biomarkers during a moderately intense exercise bout, possessing the power to support fat oxidation (Lehman *et al.* 2010).

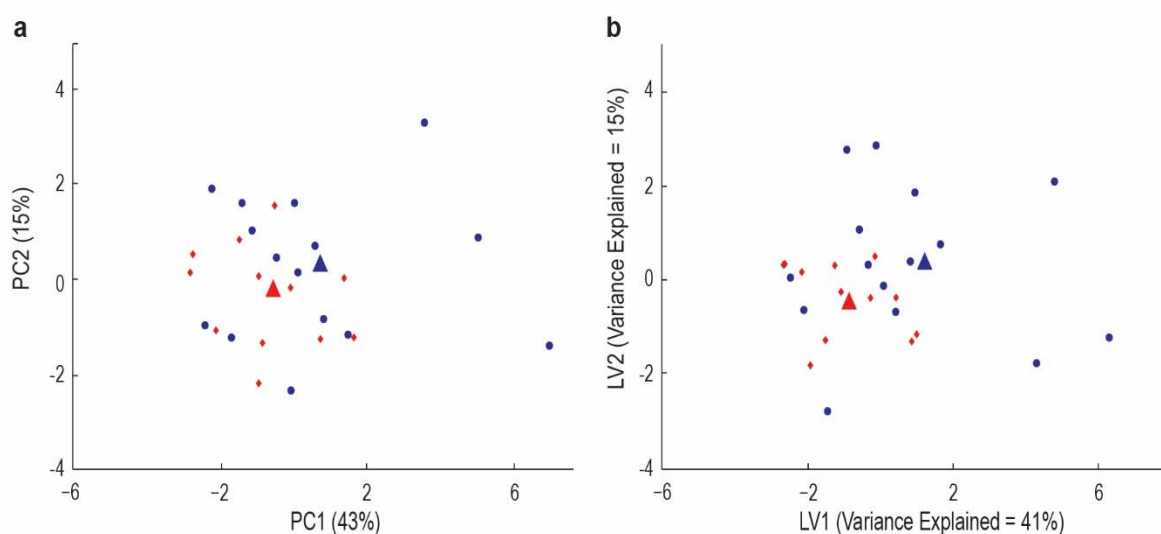


Figure 6–1 Statistical analysis of carnitines of neonates exposed (red dots) and unexposed (blue dots) to HIV-ART *in utero* as detected with untargeted LC-MS analyses with the red and blue triangles representing the center of the PCA. a: PCA score plot of normalized data of *in utero* HIV-ART-exposed infants against their unexposed controls. b: PLS-DA score plots of normalized data of *in utero* HIV-ART-exposed infants against unexposed controls.

Table 6–1 Acylcarnitines identified from the two experimental groups analyzed using multiple reaction monitoring (MRM).

HIV vs Control carnitines	Mann-Whitney test	Mann-Whitney test	Fold change	Mean	Mean	St Dev	St Dev
Acylcarnitine species	p-value	Effect size	value	HIV-ART exposed	Control	HIV-ART exposed	Control
C10 Decanoylcarnitine	0.030*	0.362	-1.327	0.034	0.045	0.010	0.020
C12 Dodecanoylcarnitine	0.037*	0.343	-1.445	0.004	0.006	0.002	0.003
C8 Octanoylcarnitine	0.068	0.286	-1.286	0.364	0.469	0.111	0.186
C0 L-Carnitine	0.082	0.268	-1.102	0.475	0.524	0.075	0.109
C6 Hexanoylcarnitine	0.165	0.188	-1.585	0.001	0.001	0.001	0.001
C14 Tetradecanoylcarnitine	0.183	0.174	-1.199	0.003	0.004	0.002	0.002
C16 Palmitoylcarnitine	0.225	0.146	-1.042	2.117	2.207	0.975	0.765
C3 Propionylcarnitine	0.225	0.146	-1.275	0.361	0.460	0.160	0.272
C4 Butyrylcarnitine	0.339	0.080	-1.152	0.032	0.036	0.015	0.032
C2 Acetylcarnitine	0.357	0.070	-1.237	2.029	2.511	1.158	2.199
C5 Isovalerylcarnitine	0.490	0.005	-1.166	0.066	0.077	0.032	0.0511

St Dev, standard deviation.

* Differences are considered to be statistically significant at a 5% level ($p < 0.05$).

Overall, the acylcarnitine profiles in cord blood from exposed and unexposed infants indicate some differentiation, but do not reflect the mitochondrial dysfunction (increased acylcarnitines) seen in the screening results reported by Kirmse *et al.* (2013). The lingering concerns about their long-term health consequences and the ramifications of the HIV-ART effects on lipid and mitochondrial metabolism nevertheless remain, given the statistically significant differences between acylcarnitines seen here in the exposed and unexposed groups.

6.2 A semi-targeted metabolomics analysis of cord blood amino acids

6.2.1 Background

The aim of the results reported in this Chapter is to describe observations on the amino acid metabolome in cord blood, derived from HIV-ART exposed and unexposed neonates. During prenatal life, the developing fetus is completely dependent on maternal physiology. The placenta is a barrier between maternal and fetal circulations and provides for the exchange of oxygen and nutrients and elimination of end products of catabolism (Latendresse *et al.* 2015). Metabolites, such as amino acids, permeate the placenta in amounts depending on their concentrations in maternal blood and on the condition of the fetus (Avagliano *et al.* 2012). Proteins do not permeate the placenta barrier significantly and are synthesized in the fetus from amino acids. Amino acids that pass from the maternal to fetal blood are thus not only a source of energy, but are mainly used for the synthesis of proteins. The results obtained from fetal development in diabetic mothers indicate that, despite optimal control of carbohydrate metabolism in the mother during pregnancy, abnormalities in amino acids metabolism are still observed. Likewise, our untargeted NMR metabolomics study indicated perturbations in amino acid concentrations due to HIV-ART exposure.

I included two cohorts of HIV-ART exposed and unexposed neonates in this study. Samples for both cohorts were prepared by me at the Academic Medical Centre in Amsterdam under the supervision of Dr Madeleine Bunders, as indicated in Chapter 1. The sample preparation included the full clinical description of all cases as well as for the HIV-positive mothers. Samples from cohort-1 were subdivided, and used by myself for the untargeted metabolomics study (Chapter 4) and the targeted experiments described in Chapters 5 and 6. Mr Nelus Schoeman, a PhD student at the Department of Analytical Biosciences, Leiden University, used the samples from cohort-2 for parallel metabolomics investigations on the effect of HIV-ART exposure on the neonate. My results on the amino acid profiles are described here, in the form

of a manuscript submitted for publication to *Metabolomics* (Moutloatse *et al.* “Metabolic risks of neonates at birth following *in utero* exposure to HIV-ART: The amino acid profile of cord blood”). The results of the metabolomics studies formed part of Schoeman’s PhD thesis (2016).

6.2.2 The amino acid profile of cord blood from HIV-ART exposed and unexposed neonates

Metabolomics (2017) 13:89
DOI 10.1007/s11306-017-1222-y



ORIGINAL ARTICLE

Metabolic risks of neonates at birth following in utero exposure to HIV-ART: The amino acid profile of cord blood

Gontse P. Moutloatse¹, Johannes C. Schoeman², Zander Lindeque¹, Mari van Reenen¹, Thomas Hankemeier², Madeleine J. Bunders^{3,4}, Carolus J. Reinecke¹

¹Centre for Human Metabolomics, Faculty of Natural Sciences, North-West University (Potchefstroom Campus), Potchefstroom, South Africa

²Department of Analytical Biosciences, Leiden Academic Center for Drug Research, Leiden University, Einsteinweg 55, 2333 CC, Leiden, The Netherlands.

³Department of Experimental Immunology, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1100AZ, Amsterdam, The Netherlands.

⁴Emma Childrens Hospital Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1100AZ, Amsterdam, The Netherlands.

Accepted: 31 May 2017; ©Springer Science+Business Media, LLC 2017

Abstract

Introduction: Untargeted metabolomics of cord blood indicated that antiretroviral therapy to HIV-infected mothers (HIV-ART) did not compromise the exposed neonates to the stress of neonatal hypoglycaemia at birth. However, energy biomarkers reflected metabolic stress and raised concern over developmental risks in some exposed newborns. *Objectives:* This study addresses the concern of HIV-ART-induced metabolic perturbations by expanding the metabolomics study to the amino acid profiles in cord blood collected at birth from newborns either exposed or unexposed to HIV-ART *in utero*. *Methods:* Amino acid profiles derived from liquid chromatographic triple quadrupole spectra of cord blood from neonates exposed and unexposed to HIV-ART (cohort-1) were investigated using a metabolomics approach. Amino acid data, generated through ultra-high-pressure liquid chromatography and tandem mass spectrometry from similar comparable cases (cohort-2), were included for comparison. *Results:* Multivariate and supporting statistics indicated differentiation between the exposed and unexposed neonates in both cohorts, caused by a general decrease of amino acid concentrations in the cord blood samples from the exposed relative to the unexposed cases. Significant up-regulation of aspartic acid in both cohorts and down-regulation of arginine, and of threonine, tryptophan and lysine in cohorts 1 and 2, respectively were observed. *Conclusions:* The benefits of ART for HIV-infected pregnant women are well established. However, the amino acid profiles of cord blood, obtained from two independent cohorts of exposed cases of their newborns, adds to observed metabolic risks with potential adverse consequences for some infants exposed *in utero* to HIV-ART, which might affect their future health.

6.2.2.1 Introduction

Antiretroviral therapy (ART) given to HIV-infected pregnant women has proved to be highly effective for the prevention of mother-to-child transmission (PMTCT) of HIV and simultaneously contributed significantly to combat the global AIDS pandemic. An untargeted NMR metabolomics study (Moutloatse *et al.* 2016) of neonates exposed *in utero* to HIV-ART, indicated that the ability of newborns to cope with transitional neonatal hypoglycaemia at birth does not seem to be markedly compromised, but indicated the presence of metabolic risk in some infants exposed to fetal ART. ART for PMTCT thus remains a matter of concern (Newell and Bunders 2013), substantiated also by investigations that have indicated dysfunctional energy metabolism and metabolic complications following prenatal ART exposure of the fetus (Blanche *et al.* 1999; Barret *et al.* 2003; Brogly *et al.* 2007; Jao and Abrams 2014).

The present study focuses on the amino acid profile in cord blood from neonates exposed *in utero* to HIV-ART. A study on the metabolic milieu of the human fetus and mother at ~20 weeks of gestation indicated that the total molar concentration of amino acids is some 2.4 times greater in fetal than in maternal plasma, with the difference in the mean umbilical concentration significantly positive for glycine, alanine, isoleucine, leucine, phenylalanine, and histidine (Soltesz *et al.* 1985). Fetal plasma amino acid concentrations obtained by cordocentesis and at caesarean section indicated that the total molar amino acid concentrations and fetal/maternal total molar concentration ratios of normal fetuses did not change significantly between the second and third trimesters (Cetin *et al.* 1990). Some recent metabolomics studies indicate perturbations in the amino acids of cord blood as a consequence of complications during pregnancies or at birth: low birth weight (Ivorra *et al.* 2012), preterm births (Tea *et al.* 2012), hypoxic ischaemic encephalopathy (HIE) (Walsh *et al.* 2012) and neonates from mothers suffering from diabetes mellitus (Hawryluk *et al.* 2016). Importantly, a review on fetal under-nutrition during pregnancy reported permanent changes in their physique and metabolism, which could relate to diseases later in life, including coronary heart disease and disorders such as stroke, diabetes and hypertension (Barker 1998). It thus seems that complexities of intrauterine fetal life are the mirror of what could happen to infant development and even health in adult life (Jao and Abrams 2014) and add a biological motivation for a study on the potential effects of *in utero* exposure to HIV-ART.

The placenta supplies and regulates the flux of amino acids in the fetal circulation (Phillips *et al.* 1978). During childbirth, the assured trans-placental nutrient and oxygen

supply is suddenly interrupted. Experiments by Kuma and co-workers (Kuma et al. 2004) with transgenic *atg5* mice, deficient in the *atg5* gene, which is essential for autophagosome formation, showed that soon after caesarean delivery the amino acid concentrations had not yet differed between the wild-type and the transgenic mice, but became significantly reduced in plasma and tissue in the transgenic mice after 10 hours in the absence of feeding. All unnourished transgenic mice died within the first day of life, prompting the authors to affirm that at birth the neonates “*encounter the first, and probably the most severe, period of starvation during their lifespan*”. These observations, signifying signs of energy depletion in the mice, imply that the effect may become compensated for through autophagy and growth factor signal transduction during the period between birth and the establishment of their ability to be nursed effectively by their mothers (Lum et al. 2005) – an observation that has become generalized to survival needs following birth in humans as well (Levine 2005).

Against this background, we speculated that the amino acid profiles in cord blood from neonates could provide further insights into the metabolic consequences for the fetus from exposure to HIV-ART *in utero*. We performed a liquid chromatography triple quad (LC-QQQ) mass spectrometry (MS/MS) targeted metabolomics analyses of cord blood samples from a cohort of HIV-uninfected but ART-exposed neonates born to HIV-1 positive mothers against a control group of HIV-ART unexposed newborns. The HIV-1 positive mothers had been diagnosed with HIV prior to their pregnancy and received antenatal combination ART (c-ART), from the first trimester of their pregnancy as part of their care (Moutloatse et al. 2016). Unpublished amino acid data from samples of a second cohort of newborns, having a similar clinical and ethnographic profile, cared for in the same clinic, were analyzed by ultra-high-pressure liquid chromatography–mass spectrometry (UPLC-MS/MS). Data from the first cohort LC-QQQ analysis were used as the primary source for investigating the effect of HIV-ART exposure on the amino acid profiles of the neonates, whereas the data from the second cohort (UPLC-MS/MS) were used for comparison. The comparison focused on the biological consequences of the HIV-ART exposure without attempting to include an analytical comparison between the two data sets. The main findings indicated some minor amino acid perturbations in the cord blood of the exposed neonates in both cohorts, which did not, however, signify critical clinical consequences at birth of exposure to HIV-ART. However, the amino acid profiles in three of the exposed cases (two from cohort-1 and one from cohort-2) gave clear indications of metabolic risks at birth for these individuals exposed to HIV-ART. The metabolomics-derived amino acid profiles, obtained from the two independent cohorts of exposed cases, caution that even small risks for

potential serious adverse effects to infants exposed to intrauterine HIV-ART warrants careful clinical attention as fetal metabolic programming *in utero* may substantially affect the future health of the child.

6.2.2.2 Materials and methods

6.2.2.2.1 Reagents and stable isotope standard

Reagents and standards for the analysis, such as acetonitrile, formic acid and methanol, were from Merk Chemical Co. (Darmstadt, Germany). Butanolic HCl (3N), acetic acid, valine, leucine, isoleucine, phenylalanine, methionine, citrulline, glycine and lysine were purchased from Sigma-Aldrich Co. The following deuterated amino acids were obtained from Cambridge Isotope Laboratories Inc. (Andover, MA, USA): [d10]-L-isoleucine, [d8]-L-valine, [d2]-glycine, [d3]-methyl-L-methionine, [d5]-ring-L-phenylalanine, [d5]-L-glutamine, [d5]-indole-L-tryptophan, [d4]-L-lysine.2HCl and [d4]-L-citrulline. The deuterated carnitine and acylcarnitine standards were obtained from Dr H. J. ten Brink, Free University Hospital (Amsterdam, the Netherlands): [methyl-d3]-L-carnitine.HCl, [d3]-acetyl-L-carnitine.HCl, [3,3,3-d3]-propionyl-L-carnitine.HCl, [d9]-isovaleryl-L-carnitine.HCl, [8,8,8-d3]-octanoyl-L-carnitine.HCl and [16,16,16-d3]-hexadecanoyl-L-carnitine.HCl.

6.2.2.3 Cases, Sample Collection and Preparation

HIV-related parameters such as viral load, CD4 count, antiretroviral medication and demographic data were extracted from records of an HIV clinic at an academic hospital affiliated to a university (see; Annexure 3.1). The exclusion criteria were HIV-exposed neonates who tested positive to any other bacterial or viral infections detected. Ethical approval for the study and for the use of the samples from the two neonatal groups was obtained from the Ethics Committee of the university that provided the samples for the investigations.

Cohort-1:

The details of cohort-1 are described in A3.1. Cohort-1 consisted of 11 *in utero* HIV-ART-exposed and 14 unexposed newborns. The mean birth weight of the infantsnewborns was 3449 (2900–4620) g. Inclusion criteria for the case study group

were gestational age at least 35 weeks [39 (37.6–41.1)] and birth weight of more than 2.5 kg (appropriate for gestational age).

Cohort-2:

The details of cohort-2 are summarized in A3.2. Cohort-2 consisted of 15 *in utero* HIV-ART exposed and 12 unexposed newborns. The mean birth weight of the newborns was 3380 (3128–3494) g. Inclusion criteria for the case study group were gestational age at least 35 weeks [40 (37.6–39.9)] and birth weight of more than 2.5 kg (appropriate for gestational age).

Controls:

The control group consisted of full-term singleton neonates born vaginally after an uneventful pregnancy and having a comparable mean birth weight with the exposed neonates from cohorts-1 and 2. The mode of delivery of infants comprising the controls was comparable with but not identical to the exposed neonates; their cord blood was sampled similarly to the procedures followed for the exposed neonates.

6.2.2.4 Experimental procedures

6.2.2.4.1 Sample extractions

Venous umbilical cord blood was obtained within 10 min of delivery, centrifuged to obtain cell-free samples and stored as soon as possible at –80 °C. The liquid chromatography triple quadrupole (LC-QQQ) method for determination of amino acids and acylcarnitines described by Rashed *et al.* (1999) and Mels *et al.* (2011) was adapted and applied to the cord blood. Fifteen microliters of cord blood was transferred to a 1.5 mL microcentrifuge tube followed by the addition of 410 µL ice-cold isotope mixture prepared in methanol. The samples were vortexed briefly and centrifuged at 6,000 rpm for 5 min. The supernatant was carefully removed and transferred to 2 mL GC-MS glass vials. The samples were evaporated to dryness under a gentle stream of nitrogen (at 65 °C).

6.2.2.4.2 LC-QQQ analysis for cohort-1

The residue of the dried samples was then derivatised with the addition of 100 µL butanolic-HCl (prepared fresh). The vials were capped and vortexed briefly before being incubated for 60 min at 60 °C and dried under nitrogen (at 65 °C). The residue

of the butylated samples was re-dissolved in 100 µL acetonitrile:water (50:50) mixture and transferred to vial inserts, then capped before LC-QQQ analysis.

LC-QQQ analysis was performed on a 1200 series Agilent LC-QQQ system linked to an Agilent 6410 MS/MS system. A C18 Zorbax SB-Aq column (150 mm × 2.1 mm × 3.5 µm) from Agilent was used for chromatographic separation of butylated compounds and kept at 30 °C throughout the analysis. Samples in the autosampler were kept at 4 °C. A one microliter sample was injected followed by the initiation of the programmed solvent gradient as discussed in A3.1 and shown in Table A3-1. The gradient for mobile phases A (0.1% formic acid in water) and B (0.1% formic acid in acetonitrile) and data on the LC-QQQ amino acid analysis are shown in Tables A3-4 and A3-5, respectively.

Agilent's MassHunter Quantitative Analysis software was used for detection and integration of the compounds' peaks according to the manufacturer's specifications. The integration of all peaks detected was checked for correctness. The integrated peak data were then exported to MS Excel where normalization (relative quantification) of the compounds was performed. The compounds which had related stable isotope internal standards were normalised with their respective isotopes. For those compounds which did not have related isotopes, isotopes that were structurally very similar and eluted closest to them were used for normalization purposes.

6.2.2.4.3 UPLS-MS/MS analysis for cohort-2

Sample collection, protein precipitation and further preparation of the samples up to the point of storage at -80°C were performed exactly as for cohort-1. The amine platform used for the LC-QQQ analysis covers amino acids and biogenic amines employing an Accq-tag derivatization strategy adapted from the protocol supplied by the Waters Corporation, and is described in detail in A3-4. An Acquity UPLC system with autosampler (Waters) was coupled online with a Xevo tandem quadrupole mass spectrometer (Waters) operated using QuanLynx data acquisition software (version 4.1; Waters). The samples were analyzed by LC-QQQ using an Accq-Tag Ultra column (Waters). The Xevo TQ was used in the positive-ion electrospray mode and all analytes were recorded by Multiple Reaction Monitoring (MRM) using nominal mass resolution.

Quality control samples

Quality control (QC) samples were used during the metabolomics analyses for quality assurance purposes for both cohorts-1 and 2. The QC pool consisted of equally pooled volumes of all study samples. A set of QC samples was then included during the

analyses of the experimental groups according to the standardized procedures used in the respective laboratories. In addition, independent duplicate samples were randomly selected where sample volume allowed and included within each batch of samples as analyzed.

6.2.2.4.4 *Statistical analysis*

Log transformation was followed to produce a more normal distribution of the variables to allow the use of parametric statistical tests. Student's t-tests, fold-change analysis and effect size were done to find metabolites that differed significantly between the groups ($P \leq 0.05$, $FC > 2$ and $d > 0.8$). Multivariate statistics was also done to visualise the natural grouping of the samples using principal component analysis (PCA). Statistical analyses were conducted with Metaboanalyst 3.0 (www.metaboanalyst.ca) and MS Excel.

6.2.2.5 **Results**

Data analyses and identification of amino acids were performed according to standard analytical and statistical procedures used for metabolomics data. The outcomes of these analyses are presented in detail in the Supplementary Information (SI, see Annexure 3). Outliers were detected for both experimental groups (Annexure A3.8) using a 95% confidence region in a Hotelling's T2 analysis and PCA, but no outliers were observed in cohort 1, cohort-2 or in the controls (Figure A3-1).

The LC-QQQ analysis of samples from the HIV-ART exposed and unexposed groups was used to identify the presence of 24 annotated amino acids (19 essential and non-essential amino acids – the leucine and isoleucine peaks were unresolved – plus ornithine, creatine, citrulline, homocysteine, gamma amino butyric and pyroglutamic acids). Five repeat analyses of the same sample using the KEMREP method with 95% bounds inclusion of QC samples (Mason *et al.* 2014) and measurement of CV values was performed for cohort 1, and all indicated high repeatability.

Unsupervised PCA and unsupervised partial least squares–discriminant analysis (PLS-DA) was used to investigate any differentiation in the amino acid data obtained from the HIV-ART exposed and unexposed among the cohort-1 neonates. First, using PCA we found that the metabolomes of exposed neonates did not group separately from those unexposed, as emphasized by the close proximity of the centroid of the 90% confidence ellipsoid shown for the two PCAs (Figure 6-2a). A similar observation

was recorded on the PCA from the exposed and unexposed groups in cohort-2 (Figure 6-2b).

Subsequently, a supervised PLS-DA was performed to maximize the discrimination between the sample groups and to identify the metabolites which potentially distinguished the groups. The PLS-DA analysis of cohort-1 somehow matched the unsupervised PCA observations. Differentiation (in terms of a greater distance between the centroids of the 90% confidence ellipsoids), but not a distinct separation, was observed between the two experimental groups (Figure 6-2c). The same holds true for the observations from cohort-2 (Figure 6-2d). The parameters of the variables important in discrimination (VIP) from the PLS-DA model were used to identify potential amino acids which best caused the partial differences in the metabolomes of cord blood from the unexposed and exposed subjects.

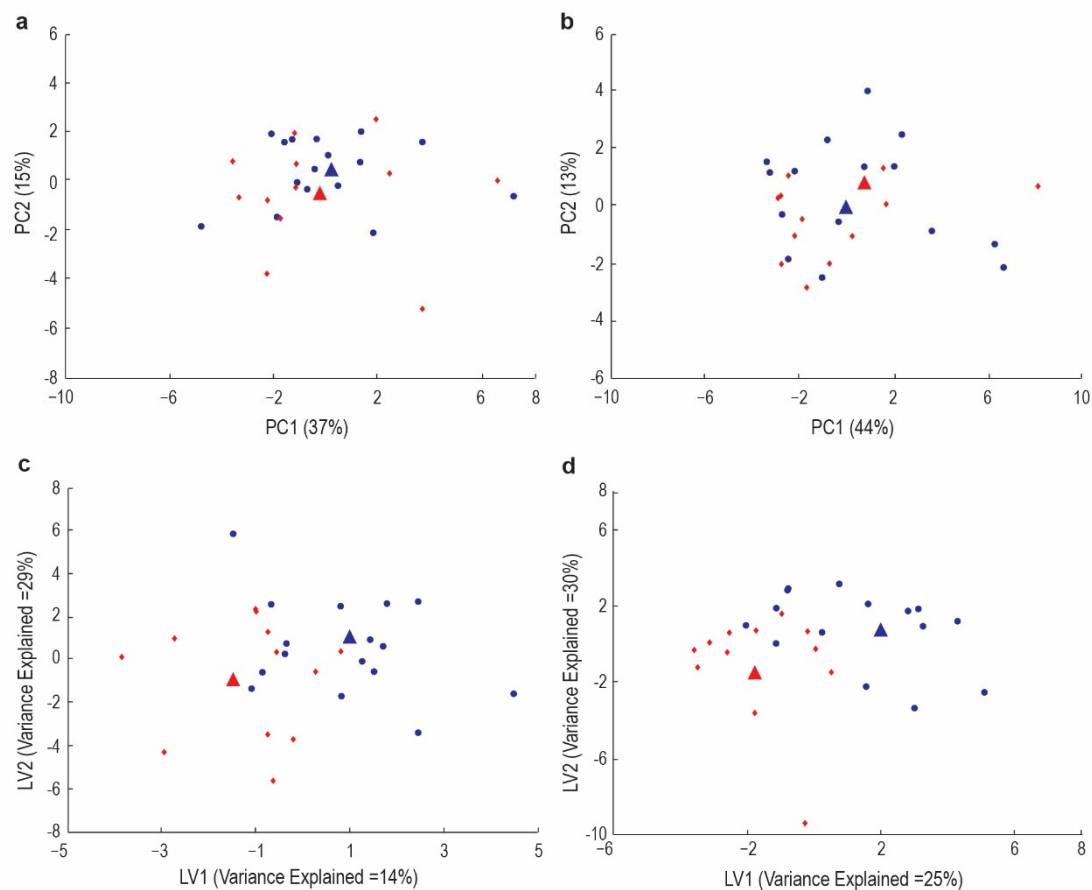


Figure 6–2: Multivariate statistical analyses to identify metabolites important in neonates exposed or not exposed to HIV-ART *in utero*. (a) PCA score plots of exposed and unexposed neonatal groups based on the quantified (μ moles amino acid) normalized amino acid data from the present experimental groups. (b) PCA score plots of the comparative exposed and unexposed neonatal groups based on semi-quantitative normalized amino acid data. (c) PLS-DA score plots of the same data as shown in a. (d) PLS-DA score plots of the same data as shown in b. Data from all HIV-exposed neonates are shown in red and for the unexposed cases in blue. The respective centroids of the 90% confidence ellipsoids are shown as red and blue triangles.

Next, a volcano plot (Figure 6-3) was executed incorporating all amino acids from our exposed and unexposed neonates. For cohort 1, aspartic acid (up-regulated in the exposed relative to the unexposed newborns: FC = -1.81) and arginine (down-regulated: FC = 1.67) were identified as the two amino acids most significantly ($p \leq 0.05$) affected by HIV-ART exposure. From the comparative study, four amino acids were indicated: aspartic acid (up-regulated in the exposed relative to the unexposed group: FC = -1.50) and threonine, tryptophan, and lysine, all three significantly ($P \leq 0.05$) down-regulated in the exposed subjects (FC = 1.34 , 1.27 , and 1.26 , respectively). Although tryptophan was down-regulated in cohort 2, the kynurenine to tryptophan ratios (data for kynurenine not included) were virtually identical for the exposed (0.00201) and unexposed (0.00205) cases, indicating that no detectable alteration in tryptophan catabolism, proposed to be observed in HIV-positive patients (Cassol *et al.* 2015), was detected due to the HIV-ART exposure.

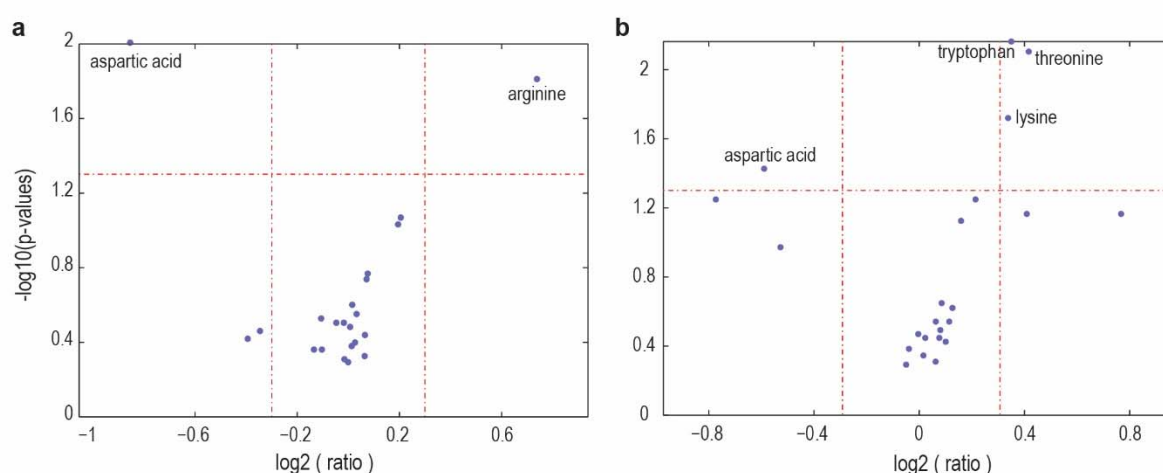


Figure 6–3: A volcano plot of all amino acids identified for both cohorts. The P -values and intensity ratios (FC) are shown as log-scaled axes, with significant differences at $P < 0.05$ (horizontal red lines) and a FC ≥ 1.20 or ≤ 0.80 (vertical red lines), indicating down-regulation or up-regulation in the exposed and the unexposed cases, respectively.

Statistical information on the quantified amino acids from cohort-1 are summarized in Table 6-2 and semi-quantified values for cohort-2 in Table 6-3. Footnotes indicated in Table 6-2 apply also to Table 6-3. According to the VIP values from the PLS-DA (Figure 6-2c and 6-2d), the slight differentiation between the exposed and unexposed groups was caused by seven amino acids in cohort 1 and by eight amino acids in cohort-2. Based on the VIP values, aspartic acid, arginine, lysine and GABA are common to both cohorts. Two other common observations are a general decrease of amino acid concentrations in the cord blood samples from the exposed relative to the unexposed neonates (60% down-regulated (FC $\leq$ 0) for cohort-1 and 80% for cohort-2); aspartic acid was the only amino acid with a low *P*-value and high VIP in common in the two groups (*P* = 0.015 and VIP = 3.24; *P* = 0.037 and VIP = 1.90, respectively).

Table 6–2: Summary of the mean concentrations and standard deviations of the amino acids with univariate and multivariate parameters for cohort-1.

Amino acid	M-W Test p-value	M-W Test Effect size	FC =	Mean CONTROL	Mean HIV	StDev CONTROL	StDev HIV	PCA Power	PLS-DA VIP
Aspartic acid	0.010	0.44	-1.81	19.21	34.71	5.06	30.76	0.15	3.24
Arginine	0.015	0.41	1.67	108.97	65.32	51.43	37.59	0.55	3.78
Glutamine	0.085	0.26	1.15	51.29	44.49	10.72	10.01	0.55	1.82
Lysine	0.093	0.25	1.14	57.96	50.63	12.03	11.47	0.55	1.68
Tryptophane	0.171	0.18	1.05	39.80	37.76	8.14	7.33	0.52	0.56
Proline	0.183	0.17	1.05	33.06	31.47	6.36	6.13	0.63	0.92
Phenylalanine	0.250	0.13	1.01	170.00	168.22	27.29	49.31	0.38	0.19
Citrulline	0.281	0.11	1.02	12.57	12.29	2.91	3.55	0.50	0.09
Asparagine	0.297	0.10	-1.08	10.71	11.53	4.53	4.69	0.36	0.36
Tyrosine	0.313	0.09	-1.01	192.75	195.09	41.89	63.53	0.45	0.42
Ornithine	0.313	0.09	-1.03	122.91	127.00	50.58	37.95	0.13	0.49
Isoleucine	0.330	0.08	1.00	574.68	571.92	145.87	169.24	0.50	0.08
Cysteine	0.347	0.07	-1.27	0.33	0.42	0.24	0.37	0.29	0.82
Histidine	0.364	0.07	1.05	493.50	471.66	141.32	128.92	0.31	1.24
GABA	0.381	0.06	-1.31	0.04	0.05	0.01	0.04	0.30	1.25
Valine	0.399	0.05	1.02	346.30	340.09	72.85	50.74	0.42	0.36
Alanine	0.417	0.04	1.01	94.54	93.69	20.91	22.75	0.50	0.97
Glycine	0.435	0.03	-1.07	144.66	155.41	35.68	49.78	0.56	0.46
Serine	0.435	0.03	-1.10	87.35	95.89	16.56	24.91	0.32	2.68
Threonine	0.472	0.01	1.05	61.37	58.69	19.45	13.05	0.19	0.17
Methionine	0.491	0.00	-1.01	24.27	24.52	5.04	5.34	0.54	0.39
Glutamic acid	0.509	0.00	-1.00	1674.94	1675.78	613.31	548.48	0.40	0.02

Significant values: *P* $\leq$ 0.05; effect size (ES) significant for the range 0.3 $\leq$ *r* $\leq$ 1.0; FC significant for TBM/control $>$ 1.0 - + for up-regulated and – for down-regulated values; PCA power values $\leq$ 0.5; PLS-DA VIP $\leq$ 1.0. Concentration units: μ mol/L

Table 6–3: Summary of the relative concentrations and standard deviations of the amino acids with univariate and multivariate parameters for cohort-2.

Amino acid	M-W Test p-value	M-W Test Effect size	FC	Mean Control	Mean HAARTs	StDev Control	StDev HAARTs	PCA Power	PLS-DA VIP
Tryptophane	0.007	0.47	1.27	7.048	5.528	1.947	0.980	0.45	2.55
Threonine	0.008	0.46	1.34	7.195	5.389	1.909	1.343	0.62	2.88
Lysine	0.019	0.40	1.26	34.286	27.126	7.899	6.017	0.47	2.54
Aspartic acid	0.037	0.34	-1.50	0.692	1.042	0.299	0.866	0.35	1.91
Glutamine	0.056	0.31	-1.71	0.335	0.572	0.464	0.497	0.32	1.09
Proline	0.056	0.31	1.16	1.729	1.490	0.576	0.457	0.55	0.87
Arginine	0.068	0.29	1.70	1.937	1.138	1.254	0.549	0.47	2.40
Methionine	0.068	0.29	1.33	1.542	1.161	0.584	0.587	0.27	1.32
Valine	0.075	0.28	1.12	0.500	0.448	0.096	0.096	0.62	0.83
GABA	0.107	0.24	-1.44	0.009	0.012	0.004	0.008	0.48	1.89
Glutamic acid	0.225	0.15	1.06	9.294	8.760	2.219	2.294	0.56	0.48
Serine	0.240	0.14	1.09	8.815	8.076	2.100	1.455	0.53	0.50
Alanine	0.287	0.11	1.08	6.839	6.318	1.817	1.630	0.50	0.50
Cysteine	0.287	0.11	1.04	0.027	0.026	0.004	0.004	0.29	0.22
Isoleucine	0.321	0.09	1.06	0.245	0.232	0.058	0.051	0.45	0.39
Leucine	0.339	0.08	-1.00	0.224	0.224	0.054	0.071	0.58	0.18
Phenylalanine	0.357	0.07	1.02	0.247	0.243	0.046	0.050	0.47	0.25
Asparagine	0.357	0.07	1.05	0.844	0.801	0.356	0.266	0.51	0.26
Glycine	0.376	0.06	1.07	10.495	9.788	3.410	2.702	0.46	0.30
Tyrosine	0.413	0.04	-1.03	3.921	4.029	0.740	0.641	0.35	0.38
Ornithine	0.451	0.02	1.01	3.408	3.370	1.084	0.878	0.58	0.22
Citrulline	0.490	0.00	1.04	0.070	0.067	0.013	0.014	0.21	0.16
Histidine	0.510	0.00	-1.03	0.062	0.064	0.011	0.018	0.24	0.90

Concentration units: $\mu\text{mol/L}$

Finally, the profiles of all amino acids for all individual cases from both cohorts were inspected. Two observations are noteworthy in the context of HIV-ART exposure. First, aspartic acid was the only amino acid observed to be up-regulated for both cohorts (Figure 6-3a and 6-3b), which is clearly emphasized by the plots shown in Figure 6-4a and 6-4b. Moreover, a single exposed subject in both experimental groups showed highly increased concentrations for aspartic acid (~3.5 times the mean value of the respective exposed group). The observations on aspartic acid therefore seem to complement each other in the two independent investigations.

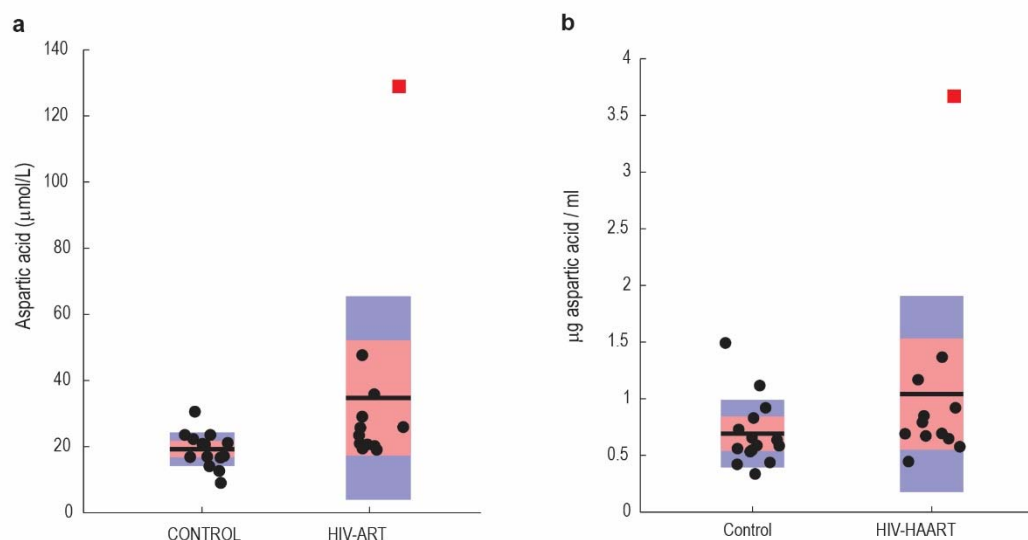


Figure 6—4: Graphs showing aspartic acid profiles from Cohort 1 (quantified concentrations) and cohort-2 (relative concentrations). Indicated in the figure are: Two infants (indicated by red squares) of HIV-ART exposure from cohort-1 (a) and cohort-2 (b) showing very high cord blood aspartic levels (~3.5 times increased) relative to the mean values of the corresponding exposed neonates. Values for all other individual cases are shown as black dots. The boxed areas represent the 95% confidence interval (orange) and 1 standard deviation (blue) about the mean (black line) [Note: HAART is comparable to ART].

Second, no substantial difference was observed for amino acids related to energy metabolism between the cases of cohort-1 and controls, except for one HIV-ART exposed subject who showed increased cord blood alanine, glycine and tyrosine (Figure 6-5a, 6-5b and 6-5c), tyrosine, methionine, proline and citrulline – a profile not shared by any other control or HIV-ART exposed case from either cohort. Additional clinical information indicated that this infant appeared to be at risk for an energy crisis at birth, indicated by low glucose (below 2.5 mM), decreased 3-hydroxybutyric acid (~20 μM) and increased hypoxanthine (~20 μM) (information included in S1).

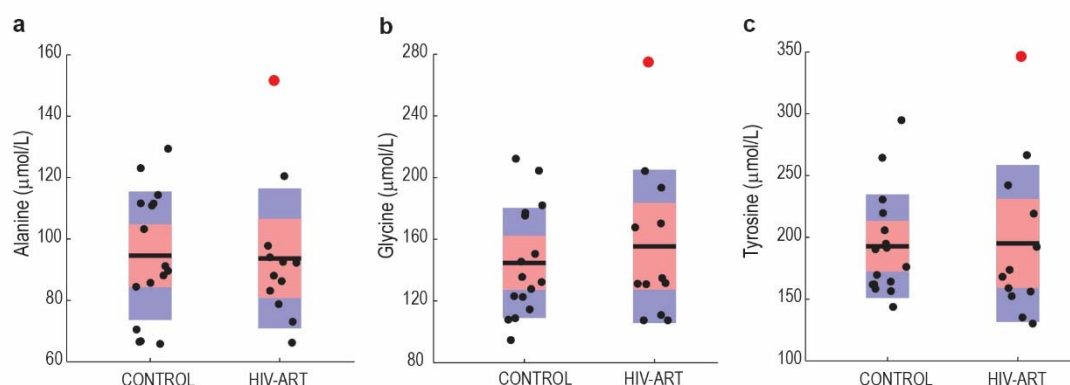


Figure 6–5: Graphs showing the quantified concentrations for alanine (a), glycine (b) and tyrosine (c) as observed in the controls and in cohort-1. A neonate (indicated by a red dot) of the HIV-ART exposure group appears to be at risk for hypoglycaemic allostasis, based on clinical parameters. Values for all other individual cases are shown as black dots. The boxed areas represent the 95% confidence interval (orange) and 1 standard deviation (blue) about the mean (black line).

6.2.2.6 Discussion

The main observations from our study are (1) that the metabolic consequences of HIV-ART exposure are already detectable at birth; (2) that a general, but not specific, decrease in cord blood amino acid values of 60–80% of the HIV-ART exposed neonates was manifested at birth; and (3) that inspection of the amino acid concentrations for all individuals revealed a few HIV-ART exposed cases presented with some significantly perturbed amino acid concentrations that calls for clinical caution.

A key question from our observations relates to the different MS platforms used to analyse the cord blood. It has been well argued (Dunn *et al.* 2012) that different workflows and analytical techniques present with limitations for the accurate comparison of metabolite profiles derived from different untargeted MS-based metabolomic studies. The comparison of experimental data (involving identical modes of analysis, accurate m/z ratios, RT-index, and fragmentation mass spectra) for each amino acid from the different studies would have been the ideal process to provide accurate comparison on the effects of HIV-ART exposure to the newborns. This was

not, however, the intention of the present study. Interestingly, however, results of relative concentrations of amino acids in cord blood present some comparative information from recent metabolomics studies on pregnancy or birth complications, as all these investigations used the highly repeatable untargeted ^1H -NMR metabolomics approach. Cohorts of low birth-weight (LBW) infants (Ivorra *et al.* 2012), those born following intrauterine growth restriction (IUGR) (Sanz-Cortés *et al.* 2013) and neonates suffering from asphyxia with possible consequences for hypoxic ischaemic encephalopathy (HIE) (Walsh *et al.* 2012; Reinke *et al.* 2013), all experienced some increased but mostly reduced amino acid concentrations. No indication of common features was observed in these studies other than dysfunctional energy metabolism – however, clinical and metabolic manifestations of mitochondrial dysfunction may vary as they depend on the tissue type affected.

Further, within a biological context, all the metabolomics studies on pregnancy and birth complications mentioned above were generally found to be complex to interpret but were regarded as important for use in research and clinical practice. Ivorra and co-workers (Ivorra *et al.* 2012) remarked: *“Even though it will be interesting and useful to clarify the mechanisms by which these metabolic changes lead to LBW consequences later in life, they are difficult to establish at present.”* We are likewise unable to give a plausible explanation of mean values of decreased metabolites observed for infants exposed and unexposed to HIV-ART. However, it has been shown that statistical information, like mean values, does not suffice in dealing with potential risks for individual HIV-ART exposed newborns (Moutloatse *et al.* 2016). The graphs shown in Figure 6-4 and 6-5 thus illustrate individual cases in terms of quantified data on the amino acids shown and invite discussion on the risks for some individual cases exposed prenatally to HIV-ART.

Firstly, significantly increased levels of aspartic acid were observed in the hippocampus of isoflurane-treated rats (Hu *et al.* 2014). It was evident that these metabolite changes were involved in the emergence of cognitive dysfunction after isoflurane anaesthesia; the researchers involved suggested that increased aspartic acid concentration in the hippocampus may be a biomarker for predicting the occurrence and monitoring the progression of cognitive dysfunction. Aspartic acid, like others, is high in the fetal brain up to the perinatal period after which it decreases towards adulthood; iso-aspartic acid, however, an isomeric form of aspartic acid, aggregates in long-lived proteins of elderly people and those suffering from Alzheimer’s disease (Ritz-Timme and Collins 2002; Shimizu *et al.* 2000). It would be speculative to link increased aspartic acid observed in our two index cases from

cohorts-1 and 2 to a potential cognitive dysfunction due to HIV-ART exposure, but the observations support the notion of risk for individuals with this prenatal history.

Secondly, one infant exposed to HIV-ART presented with reduced glucose and 3-hydroxybutyric acid but with hypoxanthine increased to a value that far exceeded 25 μ M. This concentration is considered to be an indicator of serious hypoxia (Saugstad 1975). The amino acid levels of this individual also demonstrated increased concentrations of alanine, serine and glycine (Figure 6-5), all linked to gluconeogenesis and glucogenesis at and following birth. These conditions were studied in detail by Kuma *et al.* (2004) and observed in fetal blood samples from pregnancies complicated by IUGR (Marconi *et al.* 1993). Moreover, the mechanistic target of the rapamycin complex-1 (mTORC1) pathway regulates organismal nutrient supply in response to many environmental and physiological conditions and has been shown to be necessary for neonatal autophagy and survival (Efeyan *et al.* 2013). From their studies, Sabatini and his co-workers (Zoncu *et al.* 2010; Efeyan *et al.* 2013) express the notion that a chain of events that start with the interruption of maternal nutrient supply at birth inhibits mTORC1 presumably through negative inputs from hypoglycaemia and reduced plasma amino acids. During the period between birth and first feeding, mTORC1 inhibition triggers autophagy, which generates the amino acids used to sustain plasma glucose concentrations through gluconeogenesis. We propose that our results shown in Figure 6-5a – 6-5c require a note of caution on risks for some individuals following *in utero* HIV-ART exposure.

6.2.2.7 Concluding views

The results presented in this paper support our view that metabolic changes following HIV-ART exposure are already discernible at the time of birth. Our observations continue to raise concerns regarding the safety of fetal ART exposure, despite observations that newborns exposed to this treatment appear to be healthy at birth, as (1) decreased amino acid concentrations in cord blood was seen in the two groups of newborns exposed to HIV-ART reported on here; (2) our observations add to the consequences of HIV-ART exposure reported by other investigators (Blanche *et al.* 1999; Barret *et al.* 2003; Brogly *et al.* 2007; Jao and Abrams 2014) and (3) especially by the singular perturbations in the amino acid profiles in some individual neonates exposed to HIV-ART (shown in Figures 6-4 and 6-5). In a similar vein, Jao and Abrams (2014) noted that gross measures of metabolic well-being appear to be reassuring following HIV-ART, but careful vigilance of even small risks for potentially serious

adverse effects to infants exposed to intrauterine HIV/ARVs is warranted. We recognize the outstanding and clear benefits of ART for HIV-positive pregnant women and their offspring (as recommended by the WHO). But we are convinced also that the outcomes of HIV-ART exposure discussed above fully justify continued and comprehensive investigations of the metabolic profile of newborns exposed to HIV-ART. These should be complemented by longitudinal studies on infants and children as they grow older, to gain insight on possible adverse side effects due to the exposure.

6.3 A semi-targeted metabolomics analysis of cord blood lipids

6.3.1 Background

Mr Nelus Schoeman developed the LC-MS/MS method for the determination of the lipid metabolites at the Department of Analytical Biosciences, Leiden Academic Center for Drug Research of Leiden University. He applied the method to samples from cohort 2. Schoeman indicated that cART is known to alter lipid metabolism and mitochondrial function and disturb cell homeostasis and signalling mechanisms in HIV-infected patients. The activation of these metabolic stress signalling pathways can lead to an array of reactions including activation of innate immune responses and restriction of cell growth. However, it is unknown whether *in utero* cART exposure of the fetus born to HIV-infected women leads to dysregulation of fetal metabolism with implications for growth and immunity. Using novel comprehensive targeted metabolomics analyses of cord blood plasma, he investigated the impact of *in utero* HIV and cART exposure on the metabolome of 12 infants, focusing on the lipid metabolome (lipidomics) and the mitochondrial compartment compared to 15 cases of HIV-unexposed and uninfected infants. The findings of this investigation were submitted for publication to *Nature Communications*; the abstract of the publication is incorporated in my thesis in section 6.3.2 below.

6.3.2 Lipidomics and pro-inflammatory immune responses in cART-exposed neonates

Fetal metabolic stress disrupts immune homeostasis and induces pro-inflammatory responses in HIV and cART-exposed infants

Johannes C. Schoeman¹, Gontse P. Moutloatse², Amy C. Harms¹, Rob J. Vreeken^{1‡}, Henriette J. Scherpbier³, Liesbeth Van Leeuwen⁴, Taco W. Kuijpers³, Carolus. J. Reinecke², Ruud. Berger¹, Thomas. Hankemeier^{1‡}, Madeleine J Bunders^{5,6‡*}

¹Department of Analytical Biosciences, Leiden Academic Center for Drug Research, Leiden University, Einsteinweg 55, 2333 CC, Leiden, The Netherlands.

²Centre for Human Metabolomics, Faculty of Natural Sciences, North-West University (Potchefstroom Campus), Private Bag X6001, Potchefstroom, South Africa.

³Department of Pediatric Hematology, Immunology and Infectious Diseases, Emma Children's Hospital, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1100AZ, Amsterdam, The Netherlands.

⁴Department of Gynecology and Obstetrics, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1100AZ, Amsterdam, The Netherlands.

⁵Department of Experimental Immunology, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1100AZ, Amsterdam, The Netherlands.

⁶Emma Children's Hospital, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands.

Abstract

Global roll out of antiretroviral treatment (ART) to prevent mother-to-child transmission of HIV-1 has dramatically improved survival of infants born to HIV-1-infected women. However, it is unknown why increased morbidity including enhanced susceptibility to infections and fetal growth restriction are still reported in uninfected children born to HIV-1-infected women. Viruses and/or pharmacological interventions such as ART can induce metabolic stress-signaling pathways, which skew the cell's immune response and restrict (cell) growth. We used novel comprehensive targeted metabolomics analyses of cord blood plasma, to investigate the impact of *in utero* HIV-1 and cART exposure on the infants' metabolome. Triglyceride species were increased in HIV-1-exposed uninfected (HEU) infants compared to controls, which subsequently allowed for a decrease in phospholipids. Furthermore, peroxidized lipids, indicating mitochondrial dysfunction, were increased in HEU-infants. Metabolites from compensatory metabolic pathways reducing cell stress (e.g., cytochrome p450) were reduced in HEU-infants compared to controls. The changes in the lipid and mitochondrial metabolome were associated with pro-inflammatory immune mediators including lysophospholipids and chemokines. Taken together, these data reveal that the lipid and mitochondrial metabolism is disrupted following *in utero* cART and HIV-1 exposure in the fetus leading to a pro-inflammatory milieu prior to birth with potential consequences for growth and immune development.

6.4 Conclusions

The outcomes of the three semi-targeted LC-QQQ metabolomics studies referred to in sections 6.1, 6.2 and 6.3 provided additional biochemical insight and indicated the value of the complementary analytical approach to untargeted NMR in my study. The main insights that emerged from these investigations are:

1. In the analysis of cord blood, I used LC-QQQ for the identification of the acylcarnitines and amino acids as described in Chapter 5 and in this Chapter (sections 6.1 and 6.2). I was able to obtain full quantitation of metabolites that are responsible for discriminating between the different neonatal groups in the study. This investigation supports the value of semi-targeted cord blood analysis to identify more metabolites than through NMR and thus to obtain a more comprehensive view of biological consequences of HIV-ART exposure to newborns.

2. The results from the semi-targeted metabolomics study confirmed that metabolic changes following HIV-ART exposure are already discernible at delivery, as was observed in the NMR study (Chapter 4). The information on the amino acid profiles again indicated that the neonates exposed to HIV-ART are able to cope with allostasis during transitional neonatal hypoglycaemia and do not seem as a group to be significantly compromised due to the *in utero* exposure. Our view on the benefit of ART for HIV-positive pregnant women, as recommended by the WHO, thus seems again to be confirmed, qualified by the observed risks to individual neonates exposed to HIV-ART.

3. The lipidomics study (section 6.3) covered a much wider range of metabolites than the acylcarnitines and amino acids. ART treatment for PMTCT appears to result in more comprehensive metabolite change than from the previous sections (Chapter 4, 6.2 and 6.2). In his thesis, Schoeman developed an extensive model from his results on *in utero* metabolic changes due to cART exposure and the role of the endoplasmic reticulum (ER) in stress and innate signaling. NRTI-induced mitochondrial dysfunction leads to increased levels of ROS, which are sensed by the ER. ER stress activates a cascade of responses that activate both the proinflammatory innate immune response and pathways for lipogenesis. He concludes that,

“taken together the metabolomics platforms used in [his] study provided the opportunity to assess metabolites reflecting the in vivo metabolic state in cART-exposed HEU infants and to identify pathways impacted by in utero cART exposure. Further in vitro mechanistic studies will be needed to fully understand the dysregulation of pathways that lead to the altered metabolism in cART-exposed HEU infants, and long-term follow-up studies are required to confirm our findings and determine whether these perturbations persist beyond infancy.”

This conclusion by Schoeman is a further independent finding which is in accordance with my observation from the metabolomics studies (NMR and MS), indicating to

metabolomic risks in neonates exposed prenatally to HIV-ART, as will be discussed in more detail in the next Chapter (section 7.1).

REFERENCES

SECTION 6.1: A semi-targeted metabolomics analysis of cord blood acylcarnitines

Alimenti, A., Burdge, D. R., Ogilvie, G. S., Money, D. M., & Forbes, J. C. (2003). Lactic acidemia in human immunodeficiency virus-uninfected infants exposed to perinatal antiretroviral therapy. *The Pediatric Infectious Disease Journal*, 22(9), 782–788.

Glasgow, A. M., Engel, A. G., Bier, D. M., Perry, L. W., Dickie, M., Todaro, J., et al. (1983). Hypoglycemia, hepatic dysfunction, muscle weakness, cardiomyopathy, free carnitine deficiency and long-chain acylcarnitine excess responsive to medium chain triglyceride diet. *Pediatric Research*, 17, 319–326.

Jao, J., & Abrams, E. J. (2014). Metabolic complications of in utero maternal HIV and antiretroviral exposure in HIV-exposed infants. *The Pediatric Infectious Disease Journal*, 33(7), 734–740.

Jao, J., Hazra, R., Mellins, C. A., Remien, R. H., & Abrams, E. J. (2016). Disclosing in utero HIV/ARV exposure to the HIV-exposed uninfected adolescent: is it necessary? *Journal of the International AIDS Society*, 19(1), 1–20.

Kirmse, B., Baumgart, S., & Rakhmanina, N. (2013). Metabolic and mitochondrial effects of antiretroviral drug exposure in pregnancy and postpartum: Implications for fetal and future health. In *Seminars in Fetal and Neonatal Medicine* 18(1), 48–55.

Krivitzky, L., Babikian, T., Lee, H. S., Thomas, N. H., Burk-Paull, K. L., & Batshaw, M. L. (2009). Intellectual, adaptive, and behavioral functioning in children with urea cycle disorders. *Pediatric Research*, 66(1), 96–101.

Lehmann, R., Zhao, X., Weigert, C., Simon, P., Fehrenbach, E., Fritsche, J., et al. (2010). Medium chain acylcarnitines dominate the metabolite pattern in humans under moderate intensity exercise and support lipid oxidation. *PLoS One*, 5(7), 1–12.

Lewis, W., Kohler, J. J., Hosseini, S. H., Haase, C. P., Copeland, W. C., Bienstock, R. J., et al. (2006). Antiretroviral nucleosides, deoxynucleotide carrier and mitochondrial DNA: evidence supporting the DNA pol γ hypothesis. *AIDS*, 20(5), 675–684.

Mels, C., Jansen van Rensburg, P., van der Westhuizen, F. H., Pretorius, P. J., & Erasmus, E. (2011). Increased excretion of C4-carnitine species after a therapeutic acetylsalicylic acid dose: Evidence for an inhibitory effect on short-chain fatty acid metabolism. *ISRN Pharmacology*, 2011, 1–8.

Moutloatse, G. P., Bunders, M. J., van Reenen, M., Mason, S., Kuijpers, T. W., Engelke, U. F.H. & Reinecke, C. J. (2016). Metabolic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood. *Metabolomics*, 12(175), 1–14.

Novak, M., Monkus, E. F., Chung, D., & Buch, M. (1981). Carnitine in the perinatal metabolism of lipids I. relationship between maternal and fetal plasma levels of carnitine and acylcarnitines. *Pediatrics*, 67(1), 95–100.

Poirer, M.C., Gibbons, A. T., Rugeles, M. T., Andere-Schmutz, I., & Blanche, S. (2015). Fetal consequences of maternal antiretroviral nucleoside reverse transcriptase inhibitor use in human and nonhuman primate pregnancy. *Current Opinon in Pediatrics*, 27(2), 233–239.

Rashed, M. S., Ozand, P. T., Bucknall, M. P., & Little, D. (1995). Diagnosis of inborn errors of metabolism from blood spots by acylcarnitines and amino acids profiling using automated electrospray tandem mass spectrometry. *Pediatric Research*, 38(3), 324–331.

Thorne, C., & Tookey, P. (2016). Strategies for monitoring outcomes in HIV-exposed uninfected children in the United Kingdom. *Frontiers in Immunology*, 7, 1–7.

Tsugawa, H., Tsujimoto, Y., Sugitate, K., Sakui, N., Nishiumi, S., Bamba, T., et al. (2014). Highly sensitive and selective analysis of widely targeted metabolomics using gas chromatography/triple-quadrupole mass spectrometry. *Journal of Bioscience and Bioengineering*, 117(1), 122–128.

Wilcken, B. (2010). Fatty acid oxidation disorders: outcome and long-term prognosis. *Journal of Inherited Metabolic Disease*, 33(5), 501–506.

Wilcken, B., Wiley, V., Hammond, J., & Carpenter, K. (2003). Screening newborns for inborn errors of metabolism by tandem mass spectrometry. *New England Journal of Medicine*, 348(23), 2304–2312.

SECTION 6.2: (6.2.2) The amino acid profile of cord blood from HIV-ART exposed and unexposed neonates

Avagliano, L., Garo, C., & Marconi, A. M. (2012). Placental amino acids transport in intrauterine growth restriction. *Journal of Pregnancy*, 2012, 1–6.

Barker, D. J. (1998). In utero programming of chronic disease. *Clinical Science*, 95(2), 115–128.

Barret, B., Tardieu, M., Rustin, P., Lacroix, C., Chabrol, B., Desguerre, I., et al. (2003). Persistent mitochondrial dysfunction in HIV-1-exposed but uninfected infants: clinical screening in a large prospective cohort. *AIDS*, 17(12), 1769–1785.

Blanche, S., Tardieu, M., Rustin, P., Slama, A., Barret, B., Firtion, G., et al. (1999). Persistent mitochondrial dysfunction and perinatal exposure to antiretroviral nucleoside analogues. *Lancet*, 354, 1084–1089.

Brogly, S. B., Ylitalo, N., Mofenson, L. M., Oleske, J., Van Dyke, R., Crain, M. J., et al. (2007). In utero nucleoside reverse transcriptase inhibitor exposure and signs of possible mitochondrial dysfunction in HIV-uninfected children. *AIDS*, 21(8), 929–938.

Cassol, E., Misra, V., Morgello, S., Kirk, G. D., Mehta, S. H. and Gabuzda, D. (2015). Altered monoamine and acylcarnitine metabolites in HIV-positive and HIV-negative subjects with depression. *Journal of Acquired Immune Deficiency Syndrome*. 69(1), 18–28.

Cetin, I., Corbetta, C., Sereni, L. P., Marconi, A. M., Bozzetti, P., Pardi, G., & Battaglia, F. C. (1990). Umbilical amino acid concentrations in normal and growth-retarded fetuses sampled in utero by cordocentesis. *American Journal of Obstetrics and Gynecology*, 162(1), 253–261.

Dunn, W. B., Erban, A., Weber, R. J. M., Creek, D. J., Brown, M., Breitling, R., et al. (2012). Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics. *Metabolomics*, 9(1), 44–66.

- Efeyan, A., Zoncu, R., Chang, S., Gumper, I., Snitkin, H., Wolfson, R. L., et al. (2013). Regulation of mTORC1 by the Rag GTPases is necessary for neonatal autophagy and survival. *Nature*, 493, 679–683.
- Hawryluk, J., Grafka, A., Geca, T., Dzida, G., & Łopucki, M. (2016). Study on free amino acid concentrations in umbilical cord blood plasma of newborns from normoglycaemic and gestational diabetes-complicated pregnancies. *Journal of Pre-Clinical and Clinical Research*, 10(1), 12–22.
- Hu, R., Huang, D., Tong, J., Liao, Q., Hu, Z., & Ouyang, W. (2014). Aspartic acid in the hippocampus: a biomarker for postoperative cognitive dysfunction. *Neural Regeneration Research*, 9(2), 143–152.
- Ivorra, C., García-Vicent, C., Chaves, F. J., Monleón, D., Morales, J. M., & Lurbe, E. (2012). Metabolomic profiling in blood from umbilical cords of low birth weight newborns. *Journal of Translational Medicine*, 10(1), 1–10.
- Jao, J., & Abrams, E. J. (2014). Metabolic complications of in utero maternal HIV and antiretroviral exposure in HIV-exposed infants. *The Pediatric Infectious Disease Journal*. 33(7),734–740.
- Kuma, A., Hatano, M., Matsui, M., Yamamoto, A., Nakaya, H., Yoshimori, T., et al. (2004). The role of autophagy during the early neonatal starvation period. *Nature*, 432, 1032–1036.
- Levine, B. (2005). Eating oneself and uninvited guests: autophagy-related pathways in cellular defense. *Cell*, 120(2), 159–162.
- Lum, J. J., Bauer, D. E., Kong, M., Harris, M. H., Li, C., Lindsten, T., & Thompson, C. B. (2005). Growth factor regulation of autophagy and cell survival in the absence of apoptosis. *Cell*, 120(2), 237–248.
- Marconi, A. M., Cetin, I., Davoli, E., Baggiani, A. M., Fanelli, R., Fennessey, P. V., et al. (1993). An evaluation of fetal glucogenesis in intrauterine growth-retarded pregnancies. *Metabolism*, 42(7), 860–864.
- Mason, S., Moutloatse, G. P., Marceline van Furth, A., Solomons, R., van Reenen, M., Reinecke, C., & Koekemoer, G. (2014). KEMREP: A new qualitative method for the assessment of an analyst's ability to generate a metabolomics data matrix by gas chromatography–mass spectrometry. *Current Metabolomics*, 2(1), 15–26.

Mels, C., Jansen van Rensburg, P., van der Westhuizen, F. H., Pretorius, P. J., & Erasmus, E. (2011). Increased excretion of C4-carnitine species after a therapeutic acetylsalicylic acid dose: Evidence for an inhibitory effect on short-chain fatty acid metabolism. *ISRN Pharmacology*, 2011, 1–8.

Moutloatse, G. P., Bunders, M. J., van Reenen, M., Mason, S., Kuijpers, T. W., Engelke, U. F.H. & Reinecke, C. J. (2016). Metabolic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood. *Metabolomics*, 12(175), 1–14.

Newell, M. L., & Bunders, M. J. (2013). Safety of antiretroviral drugs in pregnancy and breastfeeding for mother and child. *Current Opinion in HIV and AIDS*, 8(5), 504–510.

Philipps, A. F., Holzman, I. R., Teng, C., & Battaglia, F. C. (1978). Tissue concentrations of free amino acids in term human placentas. *American Journal of Obstetrics and Gynecology*, 131(8), 881–887.

Rashed, M. S., Ozand, P. T., Bucknall, M. P., & Little, D. (1995). Diagnosis of inborn errors of metabolism from blood spots by acylcarnitines and amino acids profiling using automated electrospray tandem mass spectrometry. *Pediatric Research*, 38(3), 324–331.

Reinke, S. N., Walsh, B. H., Boylan, G. B., Sykes, B. D., Kenny, L. C., Murray, D. M., & Broadhurst, D. I. (2013). ¹H NMR derived metabolomic profile of neonatal asphyxia in umbilical cord serum: implications for hypoxic ischemic encephalopathy. *Journal of Proteome Research*. 12(9), 4230–4239.

Ritz-Timme, S., & Collins, M. J. (2002). Racemization of aspartic acid in human proteins. *Ageing Research Reviews*, 1(1), 43–59.

Sanz-Cortés, M., Carbajo, R. J., Crispi, F., Figueras, F., Pineda-Lucena, A., & Gratacós, E. (2013). Metabolomic profile of umbilical cord blood plasma from early and late intrauterine growth restricted (IUGR) neonates with and without signs of brain vasodilation. *PloS One*, 8(12), e80121.

Saugstad, O. D. (1975). Hypoxanthine as a measurement of hypoxia. *Pediatric Research*, 9(4), 158–161.

Shimizu, T., Watanabe, A., Ogawara, M., Mori, H., & Shirasawa, T. (2000). Isoaspartate formation and neurodegeneration in Alzheimer's disease. *Archives of Biochemistry and Biophysics*, 381(2), 225–234.

Soltész, G. Harris, D. Mackenzie, IZ and Aynsley-Green A (1985). The metabolic and endocrine milieu of the human fetus and mother at 18-21 weeks of gestation. I. *Plasma amino acid concentrations*. *Pediatric Research*, 19(1),91–93.

Tea, I., Le Gall, G., Küster, A., Guignard, N., Alexandre–Gouabau, M-C., Darmaun, D. & Robins, R. J. (2012). ¹H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*, 7(1), 1–12.

Walsh, B. H., Broadhurst, D. I., Mandal, R., Wishart, D. S., Boylan, G. B., Kenny, L. C., & Murray, D. M. (2012). The metabolomic profile of umbilical cord blood in neonatal hypoxic ischaemic encephalopathy. *PLoS One*, 7(12), 1–12.

Zoncu, R., Efeyan, A. & Sabatini, D. M. (2010). mTOR: from growth signal integration to cancer, diabetes and ageing. *Nature Reviews of Molecular and Cell Biology*, 12, 21–35.

CHAPTER 7 DISCUSSION

7.1 The aim of the study

I have indicated that the introduction of ART to pregnant women to address PMTCT of HIV has become a cornerstone in public health globally and one of its greatest successes (Chapter 2). Despite the impressive reduction in HIV infection of infants following PMTCT, the treatment remains a matter of concern in view of the adverse consequences reported for some HIV-ART exposed neonates, concisely expressed by Dr Madeleine Bunders:

“The benefit of ART in the prevention of MTCT is beyond doubt, but there are reports on adverse effects on pregnancy outcome and infants currently also from high prevalence regions. Further research regarding safety is urgently required, as the number of pregnant women on ART and exposed uninfected infants is rapidly increasing.” (Chapter 1 and Newell & Bunders 2013).

- (1) Against this background, the aim of the present investigation was to address the concern regarding HIV-ART to combat MTCT, by exploring metabolic profiles in neonates exposed to ART (Chapter 2).
- (2) To achieve this aim, I chose metabolomics, a powerful, top-down systems biology approach as the analytical method (Chapter 3), to investigate the metabolome of cord blood from neonates either exposed or unexposed to HIV-ART *in utero*.
- (3) An added value of my experimental approach is that it opened up the novel opportunity to investigate whether any HIV-ART-induced metabolic perturbations are discernible at birth, a moment in life when the fetus is exposed to the trauma of birth hypoglycaemia.

To address these points, I will first discuss them from the point of view of my thesis (7.1), then regarding the views expressed by the strategic planning meeting of 2012 (7.2), and finally in terms of the present insights on NMR metabolomics of cord blood.

7.1.1 The appropriateness of metabolomics for analysis of cord blood

A metabolomics-orientated approach using umbilical cord blood has the potential to offer better understanding of neonatal homeostasis and the development or emergence of pathogenic diseases, as indicated in this thesis. This study and its

results presented here are novel and the first ever to use metabolomics to reveal the cord blood metabolome of PMTCT. The untargeted analysis of the unexplored cord metabolome and its potential application in clinical biochemistry and biomarker discovery are further described in a review presented (section 7.3) for publication. With a limited cord blood library in the Human Metabolome Database (HMDB), the use of metabolomics for this PhD investigation was able to provide new knowledge on absolute concentrations of some metabolites and contribute to the already discovered relative concentrations of others. Currently, information on the qualitative concentrations of the low molecular weight metabolites in umbilical cord blood have not been comprehensively studied and reported (Hashimoto *et al.* 2013). Therefore, a metabolomics cord blood analysis in neonatology would be highly appropriate to identify biomarkers which will better serve for disease regulation through diagnostic tools and can, one hopes, be translated into basic safe and effective clinical applications in paediatric medicine.

7.1.2 Answering the research question

The answer to the research question addressed in this study was achieved by using an untargeted NMR investigation of the cord blood metabolome of HIV-ART exposed and unexposed infants (Chapter 4). Additionally, a follow-up investigation of the results obtained from the untargeted analysis was explored using a semi-targeted LC–QQQ method. The outcomes of the untargeted analysis generated differential metabolic profiles of cord blood of HIV-ART exposed neonates against their unexposed counterparts, demonstrating that when a fetus is exposed to a stressful intrauterine environment, as observed in this study, the effects of ART and HIV on the energy metabolism (low glucose and high lactic acid and ketone bodies) are present and discernible at birth. The limitations of an NMR investigation were also revealed by this investigation, most of which were known and anticipated before the start of the experiments (elaborated on below). The observation of low levels of free carnitine and a range of amino acids in cord blood serum of HIV-ART exposed neonates prompted further analysis, which was conducted with semi-targeted LC–QQQ data. This secondary investigation revealed a differentiation of products of intermediary metabolism such as acylcarnitines and amino acids (Chapter 6). The results of the acylcarnitine analysis – decreased levels of C10-decanoylcarnitine and of C12-dodecanoylcarnitine (medium-chain fatty acids C8–C12) in cord blood of HIV-ART exposed neonates – are still only experimental and do not direct to clinical differences, although they suggest that ART exposure does alter normal fatty acid oxidation in

some infants. Keeping in mind that this is only the start of a longitudinal study. The conclusions from both untargeted and semi-targeted analysis of cord blood indicate that at birth certain newborns, delivered by HIV-positive mothers, exposed *in utero* to ART could cope with hypoglycaemic trauma at birth, although some individuals at risk were clearly indicated by the investigation (see below).

7.1.3 Views on the aim of the study, and of the contribution it makes

Regarding the recent WHO (2015) recommendations on the PMTCT programme, the effects of fetal HIV-ART exposure might be a reality for the duration of the fetus's entire developmental period even as an adult, as a consequence of the susceptible prenatal period in the fetus when its genomic identity is being established. Although the issue of prenatal genomic programming is not the topic of this thesis, the literature reports studies which have indicated the effects of maternal-fetal exposure to ART and their clinical manifestations in the infant and on genomic (epigenetic) influences. However, to my knowledge there is no existing investigation, apart from the one presented in this thesis, that has reported on the effects of mother-to-child ART exposure on the fetus *in utero* using a metabolomics approach.

The aim of this study was "to address the concern following treatment for the purpose of PMTCT by exploring metabolic profiles in exposed neonates through a metabolomics approach" – in this it was successful. Using an untargeted NMR metabolomics approach (Chapter 4), decreased levels of 3-hydroxybutyrate and hypoxanthine were observed. These identified markers are indicative of the intrauterine stress of HIV-ART exposure with consequences of a higher degree of apoptosis at birth. Although the newborns that were subject to HIV-ART exposure seem to cope with the birth stress as compared to their controls, they presented with risk metabolic factors, which warrants further monitoring. The results of the semi-targeted amino acid analysis (Chapter 6, section 6.2), such as elevated levels of aspartic acid and significantly low levels of arginine, threonine, tryptophan and lysine, confirmed that there is indeed an energy crisis in some infants born to mothers who are HIV-positive and are exposed *in utero* to ART. This is in agreement with our previous investigation, which raised concerns regarding the safety of fetal ART exposure (Chapter 6, section 6.2.2). From both these studies – the targeted and semi-targeted analysis of cord blood of uninfected HIV-ART exposed neonates born to HIV-positive mothers – it is my recommendation that a larger cohort should be recruited for further investigation. Furthermore, I suggest that the cord blood metabolome be comprehensively analysed

to develop an electronic library of cord blood metabolites which will make it less of a limiting factor in a metabolomics investigation.

7.1.4 The limitations of my study

Further limitations of my study, which were revealed by my investigation and those that were anticipated before conducting the experiments, are seen to be as follows:

- 1) The sample size of my cohort, an anticipated limitation, was relatively small (Chapter 4), although it proved to be adequate to provide a distinction between HIV-ART exposed and unexposed neonates. My experiment is one which represents a real-world situation, and so the recruitment of participants for my study was one of some difficulty. In the South African situation it still seems that the fear for stigmatization and discrimination remains an inhibiting factor for HIV positive individuals to attend conventional HIV Counselling and Testing (HCT) in health care facilities (Pérez *et al.* 2016) – a factor that led to my involvement with the HIV-ART samples collected in Amsterdam. Additionally, since my study was based on maternal-fetal interaction, a more stringent exclusion criterion for the mothers was addressed. Due to the nature of this being a metabolomics study interested in the effects *in utero* exposure of HIV-ART has on the homeostasis of the developing fetus, mothers with any confounding pregnancy complication, such as pre-eclampsia, were excluded from the study, which further decreased the number of samples available for my study.
- 2) In the background provided in Chapter 1 (Figure 1-1), I indicate that this study was part of a multidisciplinary investigation, which meant that the cord blood samples that were collected for this investigation were shared among different groups representing various disciplines. The minimum sample volume which I required for my experiments using NMR and LC-QQQ was 1 mL, in order to carry out my analytical protocol. Although there are investigations which have used smaller samples of cord blood – for example, 150 μ L by Tea *et al.* (2012), 450 μ L by Bell *et al.* (1989) and 500 μ L by Ivorra *et al.* (2012) – their NMR aims and specifications were not identical to my study. The ultimate practical aim for metabolomics studies is to have as many as possible samples of sufficient volume.
- 3) Following the untargeted run of the cord samples on the NMR spectrometer, it was discovered that there is no existing NMR cord blood

(serum or plasma) library. Although a standard serum library sufficed for the identification of my cord samples, there were unknowns which were statistically important. With the limited library available, however, they could not be identified, a handicap faced in most such NMR studies (Fan & Lane 2016; Dona *et al.* 2016). Another analytical limitation was that of the NMR spectrometer itself; despite its relatively low sensitivity, the widely used approach in NMR data processing is the binning of the signals (for our approach, see Annexure 1) or the fitting of patterns of signals to the expected metabolites. Although the latter approach is known famously for its problematic signal overlap (Markley *et al.* 2017), I was unable to fully quantify other metabolites of importance – such as methionine (see Annexure 1) – due to this very problem.

Given our well-established study design, however, these limitations had little effect on our HIV-ART investigation as we were able to differentiate the metabolomes of the cord blood of our different experimental groups

7.2 The aim defined by the international strategic planning group

I was not involved in the aims defined in 2012 by the international planning group (Figure 1, Chapter 1), and my thesis had to address the practical issue (the concern on the consequences of prenatal HIV-ART exposure for the newborn) raised by the group – the outcomes of which I discuss above. I was, however, aware of the underlying theoretical question formulated by the group: “*How is human metabolic homeostasis affected by HIV-ART exposure?*” In my study I did not focus on regulatory mechanisms to maintain biochemical homeostasis, although we frequently addressed this issue in our discussions. In fact, I included a model of perturbed metabolic homeostasis at birth in a poster that I presented at the Metabomeeting, held at the Royal Institution, London, from 10–12 September 2014. Three of our PhD students presented their results as posters at the meeting, and my contribution was the second of the series³. The poster included a proposed model on

³ Moutloatse, GP, Bunders, M., Mason, S. Engelke, UFH, Wevers, RA, Van Reenen, M and Reinecke, CJ. Explorative ¹H NMR metabolite profiling II: Differentiating cord blood profiles from controls and neonates in utero exposed to HIV and anti-retroviral treatment (ART).

metabolic homeostasis (Figure 7-1), developed in collaboration with Professor L.J. Mienie, Head of the Potchefstroom Laboratory for Inborn Errors of Metabolism (PLIEM). The model as presented here was expanded to indicate the theoretical aspects of homeostasis, allostasis and allostasis stress and will be discussed in only a cursory way.

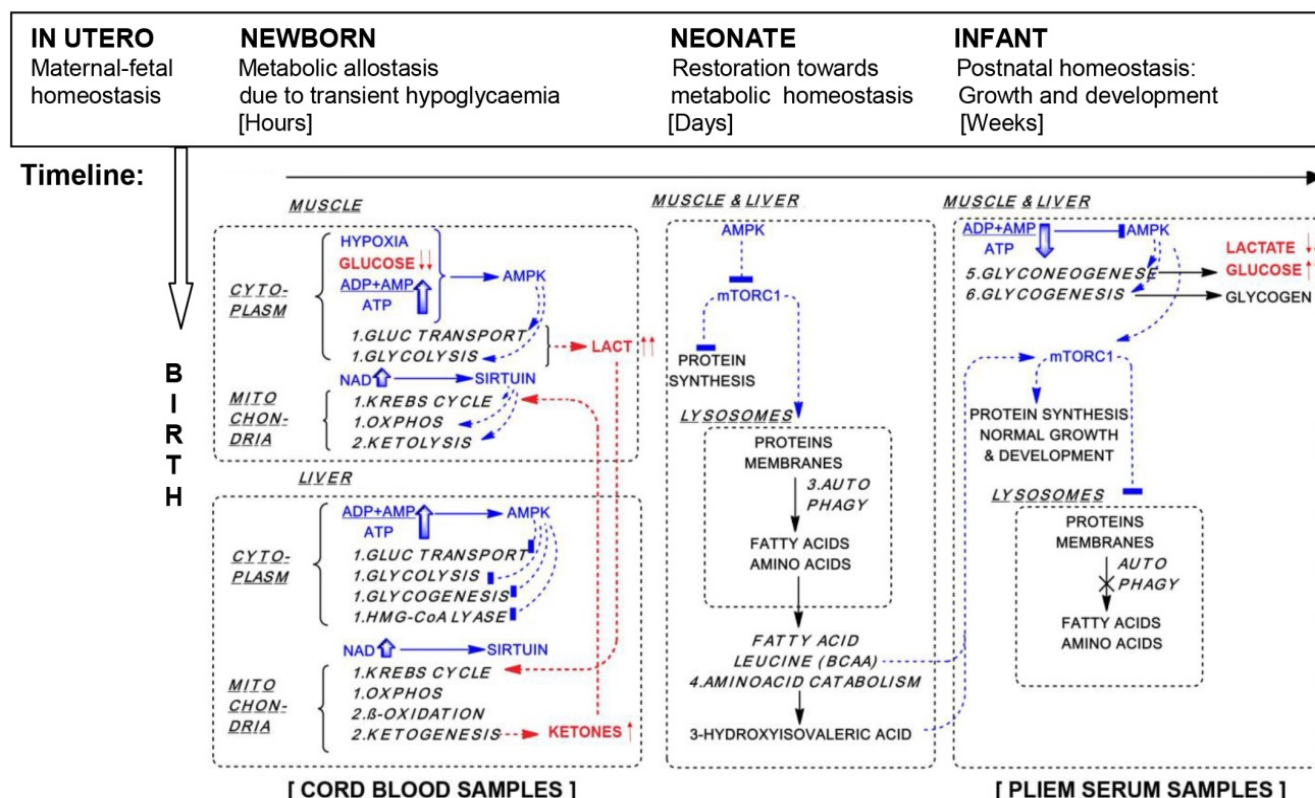


Figure 7-1: A proposed representation of perturbations of homeostasis at and following birth (Original model derived from Moutloatse *et al*, 2014 – poster presentation).³

Figure 7-1 indicates a simplified model of successive responses to restore homeostasis in the first days and weeks following birth. Key metabolites, identified in my study at the time of the conference, are indicated in red and the regulatory mechanisms are shown in blue. The main perturbation of homeostasis at birth is hypoxia and decreased glucose concentrations, resulting in transitional neonatal hypoglycaemia and hypoxia (Chapter 4, section 4.2; Moutloatse *et al*. 2016). The metabolic consequences are decreased glycolysis, decreased ATP [indicated by an increased AMP level and increased (ADP+AMP)/ATP ratio], anaerobic conditions reflected by elevated lactic acid and decreased NAD⁺. In these metabolic conditions the state of homeostasis turns to allostasis. The major molecular sensor for the AMP

level in muscle cells is AMP-activated protein kinase (AMPK), an enzyme that is activated by cellular stresses that result in ATP depletion. Mitochondrial lipolysis and ketogenesis are stimulated by the increased bodily lactic acid from the muscle tissue as well as by activation by increased NAD⁺ of the sirtuins (a family of NAD⁺-dependent histone deacetylases). Increased glucose from normal nutrition following birth leads, through the regulatory mechanisms, towards the restoration of homeostasis. Should allostasis prevail (a risk for some HIV-ART neonates as indicated in Chapters 4 and 6), a state of allostasis stress may develop with severe clinical consequences for the newborn. In healthy neonates, hours to days following birth, AMPK also inhibits mTOR, resulting in gluconeogenesis (and possibly autophagy), in parallel with inhibited amino acid consumption for protein synthesis. Finally, normal health and development ensues as observed in the metabolite profiles (increased glucose; decreased lactic acid) in serum from infants (obtained from PLIEM) that we used as controls in the study discussed in Chapter 4. A few observations made in the metabolic profiles described in Chapters 4 and 6 are not included in Figure 7-1, nor elaborated on, but may be worthwhile for further reflection and research:

- Increased myoinositol was observed in HIV-ART exposed neonates (Table 4-2: $p = 0.081$; FC = 1.87). Mammalian autophagy was shown to be regulated by the inositol signalling pathway and myoinositol is implicated in glucose homeostasis (Willams *et al.* 2006; Croze *et al.* 2013).
- A time-dependent relationship (of the order of days) between hypoxia, oxidative stress and plasma hypoxanthine (Chapter 4, section 4.6.3.2; Moutloatse *et al.* 2016) was observed in newborns (Buonocore *et al.* 2002) but not reflected in the model in Figure 7-1.
- We gave no attention to the proposed integral role proposed for leucine (isoleucine and leucine were not differentiated in our amino acid study – Chapter 6.2, Table 6-1) and its catabolite, 3-hydroxyisovaleric acid, in anabolic signalling responses – ostensibly through the mTOR-pathway (Atherton *et al.* 2010b; Eley *et al.* 2007).

Notwithstanding these comments, perturbed homeostasis, allostasis and individual risks following HIV-ART exposure have been clearly articulated through the discussion of the metabolomics results, as shown in 7.1. Greater insight into the normal cord blood metabolome is an essential aspect to address in future applications of metabolomics in neonatology, as I indicate in section 7.3 below.

7.3 Future prospects

A Mini-Review

¹H-NMR Metabolomics Studies on Neonatal Cord Blood

Submitted to Current Metabolomic

Gontse P. Moutloatse^{1*} and Carolus J. Reinecke¹

¹*Centre for Human Metabonomics, Faculty of Natural Sciences, North-West University (Potchefstroom Campus), Private Bag X6001, Potchefstroom, South Africa*

Abstract

The use of metabolomics is relatively new in neonatology. However, it has potential to provide, non-biased identification of disease-specific biomarkers in pregnancy, allowing us to better understand the pathogenesis of different neonatal diseases and provide parallel contributions in neonatal care. Despite the general increasing interest in the effect of early life status, in the metabolome of healthy neonates and from those with fetal malformations and other widespread pregnancy complications, a metabolomics approach using umbilical cord blood from neonates at birth seems to be an untapped area with unmet clinical potential in maternal-fetal medicine. Here, we review original data of published work on the current and potential use of the well-established top-down systems biology approach, with a focus only on ¹H-NMR metabolomics studies on umbilical cord blood and its recent applications in neonatology. We first developed a table with information on the cord blood metabolome, derived from metabolomics studies from plasma or serum from normal pregnancies without disease complications. Secondly, we review metabolomics studies that investigated four pregnancy complications [intrauterine growth restriction (IUGR), preterm infants, low-birth-weight neonates and neonatal asphyxia]. Next, two neonatal disease conditions are reviewed [gestational diabetes mellitus and consequences of *in utero* exposure to antiretroviral treatment given to HIV-positive mothers (HIV-ART)]. Finally, we indicate limitations in the present metabolomics studies on cord blood and give some proposals for future directions.

7.3.1 Introduction

Metabolomics – together with genomics and proteomics – has undisputedly become the most relevant conceptual driving force towards innovations in laboratory medicine. New opportunities are envisaged for clinicians, new ideas for researchers, new hopes for patients and families and new challenges for industry. Although metabolomics is still in a developmental phase in neonatology (Fanos *et al.* 2013), its potential impact was described at the 10th International Workshop on Neonatology (2014) as opening a paradigm shift towards future research and applications in the field (Fanos *et al.* 2014). Metabolomics holds high promise for gaining insight into pregnancy complications, early diagnosis of the critically ill neonate, guiding novel therapeutic choices for infants and providing scope for prognosis towards childhood development, following disease in infancy (Fanos *et al.* 2012; Katsila *et al.* 2015). Given the limitations of sampling from a neonate, proton nuclear magnetic resonance spectroscopy (^1H -NMR, further abbreviated as NMR) proved to be valuable in metabolomics investigations on maternal urine (Sachse *et al.* 2012; Diaz *et al.* 2013), and plasma/serum (Diaz *et al.* 2011), amniotic fluid (Grac $\square$ a *et al.* 2010) and plasma/serum from umbilical cord blood (Tea *et al.* 2012; Ivorra *et al.* 2012; Moutloatse *et al.* 2016). A limited number of reviews on metabolomics studies relating to pregnancy and obstetrical conditions have recently been published (Lowe *et al.* 2014; Dessi *et al.* 2015), but no review has yet appeared on metabolomics studies on the use of cord blood in normal and pregnancy-related conditions. Here, we review the published studies on NMR based cord blood metabolic profiling with the aim of highlighting key observations on the cord blood metabolome, highlight some insights that emerged from studies related to pregnancy complications and finally a few proposals on limitations and issues of importance for future NMR metabolomics research using cord blood as biofluid.

7.3.2 The cord blood metabolome

A number of comprehensive reviews have been published on the metabolome derived from various human biofluids: serum (Psychogios *et al.* 2011), urine (Bouatra *et al.* 2013), cerebrospinal fluid (Wishart *et al.* 2008) and saliva (Bertram *et al.* 2009), with partial characterization of the human metabolome in prostatic fluid (Courant *et al.* 2013) and amniotic fluid (Graça *et al.* 2007; 2009). In parallel to these exploratory investigations, the Human Metabolome Project (HMP) was initiated in the beginning of 2005 to compile the variety of information generated from high-throughput metabolomics studies and from literature surveys into one comprehensive database,

released publicly as the Human Metabolome Database or HMDB (Wishart *et al.* 2007). Studies on the composition of the human metabolome derived from various biofluids clearly indicated that the human metabolome does not constitute a unique entity comparable to the human genome. This is because the composition of the human metabolome is not solely dictated by the human genome, but includes components derived from the environment (e.g. from the diet and the atmosphere), exposure to commercial substances (e.g. food supplements and medication) and, eventually, from the gut microbiome. Thus, the human metabolome consists of the complex mixture of both endogenous and exogenous substances, per definition defined as low molecular weight substances (< 2000 Da).

Quantitative values on metabolites from cord blood (serum or plasma) are not yet covered in the HMDB, although hits related to neonatal terminology abound in the database. A search of the HMDB Version 3.6 on 24 November 2016, yielded 21,945 hits, and on “cord blood”, 1,230 for “cord plasma”, 192 for “cord serum” and 25 for “cord” with 24 for “spinal cord” and one for the “umbilical cord”, referring to hyaluronic acid (HMDB02061), the only substance described uniquely for the newborn. Hyaluronic acid “is present in high concentrations in connective tissue, such as skin, vitreous humor, cartilage, and *umbilical cord*, but the largest single reservoir is the synovial fluid (SF) of the diarthrodial joints. By contrast, the search on the HMDB for “Newborn (0-30 days old)” yielded 17,753 hits, of which many were for normal or abnormal concentrations of individual metabolites for newborns in their first 30 days of life [e.g. blood phenyl alanine = 16-71 uM – Newborn (0-30 days old) – Both sexes – Normal; blood phenylalanine = 571.0 +/- 1023.0 uM - Newborn (0-30 days old) — Both sexes – Phenylketonuria (condition)]. A number of NMR-based metabolomics studies have reported on metabolites observed in normal cord blood (plasma and serum) and in pathological conditions due to pregnancy or birth complications (Table 7-1). Relative concentrations were reported in four studies on neonatal asphyxia (Spectral intensities by Reinke *et al.* 2013), intrauterine growth restriction (IUGR); (normalized spectral intensities by Sanz-Cortés *et al.* 2013), neonates with very-low-birth-weight (VLBW) (area of normalized peak signals by Tea *et al.* 2012) and low birth weight newborns (LBW) (arbitrary units of intensity by Ivorra *et al.* 2012). Metabolites observed in the study on HIV-ART exposure were quantified for a selected chemical shift of one specific bin for each metabolite relative to an external standard (Moutloatse *et al.* 2016). Included in Table 7-1 are concentration values (relative and absolute) derived from these investigations as well as published reference values for newborns (aged 0-30 days) from the HMDB.

Table 7–1: A selection of metabolites from cord blood detected by ¹H NMR-based metabolomics.

Metabolite	Cord blood concentration Quantified Relative Mean (HIV-ART) uM Change (FC or %)	Associated pathological condition	HMDB Newborn (0-30 days)
Carbohydrate metabolism			
Glucose ^{1,3,4}	2967 (2559) μM ¹ ; Rel.(-%*)	HIV-ART ¹ ; IUGR ³ ; VLBW ⁴	1100 - 3300 μM n/r
Pyruvate ^{2,4}			
Lactic acid ¹	Rel.(FC+1.9) ² ;	Asphixia ² ; VLBW ⁴	367 – 3245 uM
Hypoxanthine ¹	Rel.(+%)	HIV-ART ¹	n/r
Ketone metabolism			
3-hydroxybutyric acid ^{1,2}	9797 (11079) uM ¹	HIV-ART ¹	
Aceton ²	23 (69) uM ¹		310 - 820 uM
Acetoacetate/acetone	318 (145) uM ¹	HIV-ART ¹ ; Asphixia ²	n/r
Protein metabolism			
Glutamine ^{1,3,4}	Rel.(FC+2.1) ²	Asphixia ²	n/r
Histidine ¹	Rel.(FC+1.5) ²	IUGR	
Leucine ^{1,2}			
Phenylalanine ^{1,3}	Rel.(+%)	HIV-ART; IUGR; VLBW	380 - 780 uM
Tyrosine ¹		HIV-ART	32 - 107 uM
Proline ⁵	397 (407) uM ¹	HIV-ART	43 - 165 uM
Alanine ⁵	Rel.(-%*)	HIV-ART; Asphixia	16 - 71 uM
Lipid metabolism			
HDL ⁴	56 (68) uM ¹	HIV-ART; IUGR; LBW	32 - 124 uM
VLDL ^{3,4}	130 (139) uM ¹	HIV-ART	111 - 329 uM
VLDL/LDL ³	Rel. ²	LBW	97 - 455 uM
Triglycerides ³	57 (62) uM ¹	VLBW	
Others			
Betaine ²	66 (67) uM ¹	IUGR; VLBW	n/r
Choline ^{2,3,5}	Rel.(-%*) ⁵	IUGR	n/r
Creatine ^{2,3}	Rel.(+%) ⁵	IUGR	n/r
Creatinine ^{1,2}			
Myo-inositol ²			
	Rel.(+%) ⁴	Asphixia ²	n/r
	Rel.(+%) ^{3,4}	Asphixia ²	24.2 – 30.8 uM
	Rel.(+%) ³	HIV-ART; Asphixia	60.8 - 107.2 uM
	Rel.(+%) ³	HIV-ART; Asphixia	29.17 - 82.22 uM
	Rel.(FC+1.2) ²		n/r
	Rel.(FC+1.7) ² ;		
	Rel.(-%*) ^{3,5}		
	64 (77) uM ¹		
	Rel.(FC+1.4) ²		
	51 (44) uM ¹		
	Rel.(FC+1.1) ²		
	241 (460) uM ¹		
	Rel.(FC+1.3) ²		

1: Moutloatse *et al.* (2016); 2: Reinke *et al.* (2013); 3: Sanz-Cortés *et al.* (2013); 4: Tea *et al.* (2012); 5: Ivorra *et al.* (2012)

The main conclusion from the results summarized in Table 7-1 is that knowledge of the cord blood metabolome is still very fragmentary. Metabolites are mostly indicated as relative concentrations and clearly indicate the need for quantified values for cord blood metabolites – but more will be needed. The fetus is attached to the maternal uterus through the placenta which is a very important organ in pregnancy, that reflects the many intricate interactions between mother, placenta and fetus. A study on preeclampsia in late pregnancy (Dunn *et al.* 2012) indicated the importance of placental tissue to define specific changes in pregnancy complications, and gave a comprehensive table of relative concentrations of metabolic features observed to be statistically significantly different ($p \neq 0.05$) in placental tissue from uncomplicated and preeclamptic pregnancies. Dunn and co-workers (Dunn *et al.* 2012) suggest that changes observed in placental tissue should be integrated with data from biofluids including maternal serum/plasma, urine, fetal blood and cord blood to provide for a holistic view on these interactions and for insights on pregnancy complications.

7.3.3 Metabolomics insights on pregnancy complications

NMR-based metabolomics investigations have offered valuable insights on many diseases and most recently on pregnancy outcomes and neonatal pathologies. Presented in this section are exemplary studies that have reported in literature on clinical metabolomics in neonatology using cord blood and NMR metabolomics.

7.3.3.1 Intra uterine growth restriction

IUGR also referred to as intrauterine growth retardation (Dessi *et al.* 2011), is defined as the inability of the fetus to reach its full genetic growth potential, often characterized by a low fetal birth weight and body mass index lower than normal for its gestational age (Committee on Practice Bulletins Gynecology 2011) Intrauterine fetal growth relies on a balance between the maternal-placental nutrient supply and the genetically established growth potential of a fetus (Battaglia & Meschia 1986). Pregnancies complicated by IUGR are associated with increased risks for adverse perinatal outcomes like a damaged neuroendocrine system, cardiovascular dysfunction and deformations, but also consequences related to dysfunctional metabolism such as birth hypoxia and metabolic syndromes which have implications for early childhood and later in adult life (Dessi *et al.* 2015).

Although IUGR affects 7–10% of all pregnancies (Bernstein *et al.* 2000; Diderholm 2009), there is currently little knowledge regarding the overall metabolic status of

neonates with the condition. In a metabolomics investigation of umbilical cord blood plasma from early- and late-IUGR neonates, Sanz-Cortés *et al.* 2013) used NMR spectroscopy to characterize a full spectrum of metabolic changes in umbilical vein plasma of newborns diagnosed with varying forms of IUGR. In their investigation, a prospective cohort study, they included neonates who were clinically clearly diagnosed in terms of umbilical artery Doppler pulsate index (PI) as IUGR (n=79) compared with matched controls (n=79). Their patient group was divided into two sections, those diagnosed with early IUGR (n=23) and those who were diagnosed with late IUGR (n=56) against a control group of singleton neonates who presented with birth weights appropriate for their gestational age (AGA).

At delivery, 500µl of umbilical vein blood after cord clamping was collected in EDTA-treated tubes. Samples were analysed on an NMR with their spectra recorded on a Bruker Ultrashield Plus Avance II 600 MHz spectrometer. The metabolomics data matrix generated consisted of scaled data for the unsupervised partial least squares (PCA), used to observe the clustering or separation of the experimental groups. Further data analysis was conducted using supervised multivariate partial least squares discriminant analysis (PLS-DA) and orthogonal PLS-DA (OPLS-DA). The PCA score plot results corresponding to early-onset IUGR against matched AGA ($R^2 = 0.63$; $Q^2 = 0.37$) revealed a reasonable separation. NMR regions responsible for separation were identified as those characteristic of lipids, lactate and methyl groups of amino acids such as valine, leucine, isoleucine and threonine; glucose was a strong indicator as the relevant metabolite mainly responsible of the early-IUGR case and controls. An improved separation was observed in the OPLS-DA model ($R^2Y = 0.79$; $Q^2 = 0.58$), revealing very high levels belonging to lipid molecules [unsaturated and very low density lipoprotein (VLDL)] both in early- and late-onset IUGR cases compared to controls. Some informative metabolic observations in the IUGR study were: (1) triglycerides were found to be increased up to 33% in early-onset IUGR cases when compared to controls, which was consistent with previously published evidence. This finding suggests that triglycerides might be early predictors of coronary arterial disease, known to manifest as cardiovascular disease later in adult of IUGR patients (Kaser *et al.* 2001). (2) Increased unsaturated lipids and VLDL levels occurred in samples from both early and late IUGR. The authors indicated that these results agree with previous reports which reflect altered, pro-atherogenic lipid and cholesterol metabolism in IUGR fetuses (Pecks *et al.* 2012), but could also be explained by chronic prenatal exposure to hypoxia that can induce significant hyperlipidaemia during perinatal life (Raff *et al.* 2000). (3) Regarding amino acids and derivatives, early IUGR

showed significantly increased glutamine and creatine levels, whereas the amounts of phenylalanine and tyrosine were decreased in samples from early- and late-IUGR. Valine and leucine were decreased in late IUGR samples. A decrease in glucose for early IUGR cases. Valine and leucine were decreased in late-IUGR sample. A decrease in glucose for early-IUGR relative to controls as well as choline and an even lower phenylalanine level were identified. Decreased glucose, phenylalanine and choline were associated with adverse neurological outcomes (Montoya *et al.* 2000).

The strength of the IUGR study were the well-defined cohorts that were constructed prospectively and the controls which were selected to match the affected neonates through gestational age to avoid potential confounding influences; the authors asserted that the findings from the metabolomics study were in agreement with their current clinical practice. However, they concluded that IUGR at this stage is not associated with a unique metabolic profile, but that important differences were observed which may be already be of use in both clinical practices and in further research.

7.3.3.2 Preterm infants

A variety of common pregnancy complications, such as IUGR mentioned above, is owed to placental dysfunction. During the final stages of pregnancy, the human fetus triples in body weight and relies heavily on the maternal-fetal transfer of nutrients to cover the latter's energy needs (Platt & Deshpande 2005). An NMR metabolomics study (Tea *et al.* 2012) probed to link between metabolic adaptations and premature birth whilst assessing the metabolic profile of VLBW neonates and their mothers. The researchers sought insight into the maternal-fetal metabolite exchange without the influence of a distributing bias. Premature births, resulting in VLBW infants, are classified as the delivery of a baby before 33 weeks of pregnancy (Alexandre-Gouabau *et al.* 2013). In an earlier study Bell *et al.* (1989) used targeted NMR to examine plasma obtained from umbilical and maternal veins, the authors describe findings on lactate, alanine, valine, and lipoproteins in both maternal and venous blood in mothers who delivered healthy normo-ponderal full term (FT) infants. Similarly, the recent NMR metabolomics study (Tea *et al.* 2012) but with the inclusion of VLBW neonates and the different modes of delivery to expose the general differences between term and preterm deliveries. They reported on well-characterized neonates with VLBW (n=8) against a control group of healthy FT (n=8), normal-delivery mother-infant pairs. Samples of arterial and venous umbilical cord blood were collected directly after birth in EDTA-coated vacutainers. Similarly, maternal venous blood for each mother-infant pair was also collected at delivery to observe differentiation between preterm deliveries

of VLBW neonates against FT controls. After minimal sample preparations, the samples were run on a Bruker Avance 500 MHz spectrometer, equipped with a cryogenic probe where all high resolution blood plasma spectra were recorded. An unsupervised PCA and an additional supervised PLS-DA were performed. The NMR results displayed the possible discovery of 26 metabolites, indicating the high-throughput capabilities of NMR as a detection tool in neonatal research. The results uncovered in this investigation revealed that a clear-cut metabolic difference between umbilical arterial, umbilical venous and maternal plasma exists, with the term of delivery playing a definitive role. The metabolites identified were typically low-molecular-weight compounds such as amino acids, glycolysis metabolites, ketone bodies along with lipoproteins, unsaturated lipids, glycoproteins and albumin-lysyl which are the larger molecular mass metabolites. There was no noticeable spectral difference between arterial and venous cord plasma; however, due to the presence of protein lipid resonance, which is rarely seen in cord plasma and maternal plasma was observed.

In contrast to mothers delivering at full term, those giving birth to VLBW infants presented with biological perturbations indicate by higher levels of lipids, pyruvate, glutamine, valine and threonine with significantly lower levels of acetate and isoleucine. These results were indicative of the metabolic effect of preterm deliveries on maternal plasma. In terms of the effect of gestational age on both venous and arterial cord plasma of FT and VLBW neonates, supervised multivariate analysis did not indicate a significant metabolic signature that related to the term of delivery. However, in venous cord plasma of VLBW preterm deliveries, univariate analysis revealed significant decreases in levels of lipids, pyruvate, glucose, acetone, albumin-lysyl and noticeable increases in levels of formic acid and glutamine/glutamate, which were also characteristic of arterial cord plasma. The decrease in the levels of lysyl groups of albumin in VLBW is described to be resulting from a certain degree of oxidative stress (Rodiño-Janeiro *et al.* 2010). This oxidative stress marker was also identified in cases of VLBW maternal-fetal pairs irrespective of mode of delivery, a relevant discovery since half of VLBW neonates suffered from hypertension. In normal pregnancies, however, a lower level of lysyl-albumin is associated to signs of hypoxia and oxidative stress (Prefumo *et al.* 2007).

The NMR metabolomics study of cord blood (Tea *et al.* 2012) revealed new insights on the potential differences in maternal-fetal nutrient exchange and successfully established a detailed metabolic profile from maternal as well as from venous and arterial cord blood plasma, indicating a difference in metabolites based on term and

mode of delivery. The authors concluded that the simple, non-invasive mode of collection of cord plasma is particularly suited for metabolomic studies.

7.3.3.3 Low birth weight neonates

Developmental plasticity is “the ability of a single genotype to produce more than one alternative form of structure, physiological state or behavior in response to environmental conditions” (Barker 2004). This concept suggests that early life conditions can modulate chronic disease risks in adulthood (Burdge & Lillycrop 2010). Infants born with LBW have been known to experience accelerated postnatal weight gain, which is associated with increased risks of diabetes, obesity, hypertension (Gluckman *et al.* 2008) and cardiovascular diseases observed in both human (Barker *et al.* 2005) and animal models (Isganaitis *et al.* 2009). An NMR metabolomics study (Ivorra *et al.* 2012) addressed the knowledge gap concerning the available data identifying a metabolite profile in umbilical cord from LBW children. Against this background their objective in this metabolomics investigation was to “analyze the metabolomic profile at birth of LBW neonates and determine if the metabolomic profile of plasma from their umbilical cord blood differs as compared to normal weight newborns”, whilst concurrently analyzing the metabolomic profiles of the mothers at the time of birth. For this purpose, newborns born at term (by vaginal or caesarean) with gestational age ≥ 37 weeks, in an uneventful pregnancy were recruited. Neonates with a birth weight between the 75th and 90th percentile were regarded as controls ($n = 30$), whereas the patient group ($n=30$), included LBW neonates whose weight was lower than the 10th percentile ($< 2.5\text{kg}$). Clinical information of mothers displayed no differences in maternal age, maternal BMI, pregnancy weight gain or mode of delivery between the two groups. The infants however displayed no differences in maternal age, maternal BMI, pregnancy weight gain or mode of delivery between the two groups. The infants however, displayed differences in blood pressure, with the LBW group demonstrating significantly lower systolic values than the control group. Umbilical venous cord blood was collected at birth immediately into EDTA-treated tubes. Similarly, maternal peripheral blood was collected but 2-4h after delivery in non-fasting conditions. Following nominal sample preparation, 500 μl of each sample was analysed using NMR Bruker Avance 600 MHz NMR spectrometer system. All spectra were processed using Topspin 1.3 software and in-house scripts were applied for data analysis. From the loading plots of their spectral data the authors were able to identify 34 variables, whereas multivariate analysis identified 17 variables as being the significant metabolites responsible for the observed separation. From the semi-quantified metabolites, proline, free choline, glutamine, alanine and glucose were

found to be significantly lower in LBW neonates than their controls, while phenylalanine and citrulline indicated significant increases between the LBW infants and control group. A Pearson correlation analysis was performed on each of these metabolites. Interestingly, a positive correlation was observed between birth weight and free choline ($r = 0.50$), proline ($r = 0.29$) and glutamine ($r = 0.36$). Additionally, an inverse association was observed between citrulline (-0.40) and phenylalanine ($r = -0.34$) and birth weight. Moreover, the observation of increased phenylalanine and unbranched amino acids, in LBW was reported to be associated with insulin resistance (Newgard *et al.* 2009) and development of type 2 diabetes (Wang *et al.* 2011).

Free choline levels were identified to show the biggest difference in this study, with decreased levels in LBW group. Choline is a known donor of methyl groups through its metabolite betain for the production of S-adenosylmethionine (SAM), a substrate of DNA and histone methyltransferases (Mehedint *et al.* 2010). The researchers gave a novel interpretation from this observation by speculating that maternal choline alterations during pregnancy modify fetal histone through DNA methylation – observed in experimental animals (Davison *et al.* 2009) – suggesting that a concerted epigenomic mechanism contributes to the long-term developmental effects of varied choline intake *in utero*.

This study is one of the first LBW investigations to use cord blood metabolomics in order to identify metabolites that could be related to the abnormalities associated with the long-term development of the increased cardio-metabolic risk linked to LBW and its consequences later in life. However, the researchers conclude that the meaning and the potential value of their findings on biomarkers of risk in LBW neonates need to be addressed in future studies.

7.3.3.4 Neonatal asphyxia

The stress of labour is known to pose many challenges for the emerging infant (Platt & Deshpande 2005). Perinatal asphyxia (PA), identified as neonatal oxygen deprivation during birth, is a major cause of neonatal death with an incidence of 2 per 1000 deliveries (Lawn *et al.* 2005).

A recent review covered metabolomics research on perinatal asphyxia condition, examining in detail the studies reported both on animal and human models, conducted on various metabolomics platform (Fattuoni *et al.* 2015). In their NMR metabolomics study, Reinke and co-workers (2013), reported on the consequences of perinatal asphyxia. Most infants completely recover completely from hypoxia linked to asphyxia, but the researchers emphasize the relevance of their study as some neonates do not

recover from the hypoxia and develop hypoxic ischemic encephalopathy (HIE) with lasting neurological sequelae such as cerebral palsy, seizure disorders, cognitive delays, and motor disabilities. HIE remains a global health concern, being responsible for 1 million full-term neonatal deaths globally each year (Lawn *et al.* 2005). A prime motivation for their study was to develop a more inclusive metabolic profile of perinatal asphyxiation and to determine its association with neurological injury; considerable attention was given to the stratification of the cohorts of affected and control neonates. Infants were included only if over a gestational age of 36 weeks. Fifty-nine at-risk neonates were enrolled, together with 59 1:1 matched healthy controls, and classified according to disease severity (n = 25, HIE; n = 34, non-HIE PA).

Proteins were removed from 200 μ L of each serum sample using a commercial kit for perchloric acid precipitations of proteins, removed by centrifugation followed by neutralizing of the pH. In their method, they used a 600MHz NMR spectrometer to derive umbilical cord serum metabolites. Their untargeted metabolic analysis detected 37 quantified and reproducible acid-soluble metabolites: organic acids central to energy metabolism, amino acids and their catabolic intermediates, urea cycle metabolites, ketones, and phospholipid groups. Using multivariate analysis 18 metabolites between asphyxia infants and their controls were identified as significant, whereas only 13 metabolites were significantly different between the HIE infants and their matched controls.

The study also revealed a metabolite profile with an association of both asphyxia and HIA, which included metabolites such as alanine, choline, creatine, glycerol, isoleucine, lactate, leucine, myo-inositol, pyruvate, phenylalanine, succinate and valine, which were increased in both asphyxia and HIE cases. Methionine was only uniquely elevated in HIE; similarly acetone, betaine, creatinine, glucose, 3-hydroxybutyrate and O-phosphocholine were uniquely identified as elevated in asphyxia cases. Furthermore, through multiple linear regression (adjusted $R^2 = 0.4$), 3-hydroxybutyrate, succinate, glycerol and O-phosphocholine were observed to be the most predictive of severe HIE.

Notably, in a previous study from the same research group, Walsh *et al.* (2012) used the same cohort for an LC-MS/MS investigation. In that study, neonates suffering from asphyxia could be classified as HIE or non-HIE from the metabolite profiles. Three metabolite groups were noted as important in this regard: amino acids, acylcarnitines and phospholipids. However, the researchers were unable to detect central energy metabolites in the LC-MS/MS study, which they observed in their NMR investigations (Reinke *et al.* 2013), which gave insight to primary energy metabolism variations

arising from neonatal asphyxiation and its implication for neurological injury. Evidently, their finding of depleted ketones (acetone, 3-hydroxybutyrate) in severely affected HIE infants is described as a result of hypoxia, leading the infants to be more susceptible to hypoxic injury. In contrast acetone and 3-hydroxybutyrate levels were elevated in asphyxiated infants, which suggests that unlike in HIE, systemic metabolic responses such as ketogenesis play a large role in preventing neurological injury during prenatal asphyxia.

Data from these related studies revealed that cord serum metabolites is altered significantly in perinatal asphyxia but more importantly that cord blood can be used for metabolomic analysis for detecting encephalopathy in the neonate in its first stages of life. These studies affirm that through the more inclusive metabolomic profile of perinatal asphyxiation and its association with neurological injury (HIE), both protective and pathological consequences of perinatal asphyxiation can be revealed and provide insight into mechanisms underlying HIE. These comprehensive studies on the same cohort of cases and controls emphasize the added value of complementary metabolomics approaches in investigating complex metabolic perturbations.

7.3.4 Fetal-maternal disease conditions

7.3.4.1 Gestational diabetes mellitus

Pregnant women who have never had diabetes before but present with high blood glucose levels during pregnancy, are clinically defined as having gestational diabetes. This condition is characterised by carbohydrate intolerance, first recognized during pregnancy (McCance 2011). Infants of diabetic mothers (IDM) and of mothers with gestational diabetes (IGDM) are known to encounter risks of perinatal mortality and morbidity (Hay 2012). These risks include macrosomia, neonatal hypoglycemia, respiratory distress syndrome and consequences faced later in their adult life such as type 2 diabetes mellitus, childhood obesity and metabolic syndrome (Dessi *et al.* 2015). In a NMR metabolomics study, Dani *et al.* (2014) based their work on the hypothesis that the metabolomics profile of FT neonates born to gestational diabetic mothers would be different from infants born to healthy mothers despite the strict control of the mother's gestational diabetes. In this investigation 5mL of umbilical arterial blood was collected from 30 term (≥ 37 week) IGDMs and 40 infants from healthy mothers, who served as controls. The metabolomic analysis of these samples was conducted on a Bruker 600-MHz NMR spectrometer. Multivariate analysis PCA was used to derive a

general overview of the data and OPLS-DA to identify discriminating metabolites between the two sample groups. Although the results obtained from the PCA gave no evidence of relevant separation between the two groups, the OPLS-DA offered good recognition accuracy between them. From the metabolites identified glucose was significantly lower in IGDM neonates than their controls whereas pyruvate, histidine, alanine, valine, methionine, arginine, lysine, hypoxanthine, α -ketoisovaleric acid, lipoprotein and lipid metabolites were observed to be increased but showed no clinical differences. Lower levels of serum glucose in IGDMs than their respective controls have been associated with neonatal hypoglycemia caused by the sudden decrease in maternal glucose supply at birth combined with neonatal hyperinsulinemia. The researchers indicate that a limitation in their study is the lack of follow-up of metabolomics profile to evaluate the duration of differences between IGDMs and controls. However, a metabolic profile of term IGDMs at birth was discovered, which differentiated from normal-term infants born to regular pregnancies. The conclusion was that prolonged fetal exposure to hyperglycaemia during gestation can alter the neonates' metabolic profile at birth. The adverse effects of GDM pregnancies in neonatal medicine therefore still calls for early and complete biomarkers for the prediction and screening of the onset of GDM.

A systemic review of metabolite profiling in gestational diabetes mellitus (Huynh *et al.* 2014) summarized the extensive search of biomarkers in research conducted over the last decade using a metabolomics approach to profile the metabolome of mothers (in terms of blood serum, plasma and urine) and their infants (by monitoring amniotic fluid, umbilical cord blood, meconium and urine) in order to identify new knowledge and insights regarding the pathology of GDM. A recent holistic NMR metabolomics investigation delineated the maternal blood metabolome related to postdiagnosis GDM subsequently envisaged to identify biomarkers of both pre- and postdiagnosis GDM of maternal plasma and corresponding lipid extracts (Pinto *et al.* 2015). These results showed that pre-diagnosis GDM stages (2–21 gestational weeks prior to diagnosis) involve early metabolic variations, detectable in both maternal plasma and corresponding lipid extracts. Included in their paper, some considerations were drawn in relation to possible future applications of NMR metabolomics in clinical GDM management. In addition, Dani and co-workers (2014) noted that despite the great potential of prenatal urine and blood for the use in novel, less-invasive diagnosis and monitoring methods, the exploitation of metabolite profiling of these biofluids is still insufficiently developed, but confirmed that their results suggested that prolonged fetal exposure to hyperglycemia during pregnancy can change an infant's metabolomic

profile at birth, underscoring the value of metabolomics studies in the field of neonatology.

7.3.4.2 HIV-ART exposure

HIV is globally recognized to be responsible for a large proportion of deaths in countries with high HIV prevalence, including a substantial burden of HIV-caused maternal deaths (134 000 deaths annually world-wide with 84 000 in sub-Saharan Africa) (Say *et al.* 2014). Antiretroviral therapy (ART) is effective in reducing maternal mortality among HIV-infected women (Liotta *et al.* 2013), and its success in preventing mother-to-child transmission (PMTCT) of HIV, the cause of a global paediatric pandemic, is widely recognized in public health and one of its greatest successes (Chasela *et al.* 2010). ART treatment for HIV positive pregnant women (HIV-ART) remains a matter of concern (Newell & Bunders 2013) on the basis of several reports on dysfunctional energy metabolism and metabolic complications following prenatal ART exposure to the fetus (Blanche *et al.* 1999; Barret *et al.* 2003; Brogly *et al.* 2007; Jao & Abrams 2014).

NMR analysis of cord blood from 11 clinically well-defined *in utero* HIV-ART-exposed and 14 unexposed newborns, as well as serum from 8 control infants, were used to generate spectral data for a metabolomics study to address the concern over HIV-ART exposure *in utero* (Moutloatse *et al.* 2016). High-resolution NMR spectra of ~700 µl deproteinized cord blood serum samples at pH=2.50 were recorded on a Bruker 500 MHz Bruker NMR spectrometer, using Bruker Topspin (V3.1) and Bruker AMIX (V3.9.12) software for binning and quantification (Engelke *et al.* 2006; Engelke *et al.* 2008; Wevers *et al.* 1995). Division into bins across the spectral regions was performed automatically and variable-sizing was used to create a data matrix of 114 bins/variables. The score plots from unsupervised PCA based on data from the infant and unexposed neonatal groups from the normalized bins indicated a complete separation between the infant controls and the unexposed neonates. Subsequent PCA loadings S-plots and PLS-DA revealed that 11 metabolites provided for qualitative information on the hypoxic and hypoglycaemic state prevailing at birth with carnitine, glucose and lactic acid being the three most significant markers. Decreased free carnitine in the neonates ($p = 0.0001$) could reflect a reduction in intrauterine maternal-fetal transfer in late gestation (Novak *et al.* 1981) while depletion of glycogen was implicated in the significant low levels of glucose in the neonates ($p = 0.0001$) presenting with hypoglycaemia as the most common manifestation at, and following, birth. Increased levels of lactic acid ($p = 0.0001$) and ketone bodies (3-hydroxybutyric

acid - $p = 0.0015$) indicate their participation as additional key energy resources for the critical energy dependency of the newborn brain.

Against these general observations of the cord blood profiles from the exposed, unexposed and infant controls, two important observations were made from further analyses of the metabolomics results: Firstly, calculation of the potential ATP generation from residual glucose and increased lactic and 3-hydroxybutyric acids (derived from lipolysis at birth) shows that the very low glucose levels were fully compensated for by the alternative metabolic fuel provided by increased lactic acid, and mobilized 3-hydroxybutyric acid. The key conclusion is that the metabolic response towards hypoglycaemia at birth for the neonates exposed to HIV-ART was comparable with that of the unexposed neonates. The response of the HIV-ART-exposed neonates to allostasis at birth did not seem to be compromised due to the *in utero* environment. Second, two metabolites [3-hydroxybutyric acid (odds ratio: 1.69 (95 % CI 1.08–3.15))] and hypoxanthine (odds ratio: 2.01 (95 % CI 1.14–5.130))] were observed to differentiate significantly between HIV-ART-unexposed infants and some *individuals* who had exposed. Moreover, cord blood levels of hypoxanthine greater than 25 μM (Saugstad *et al.* 1975) demonstrated a potentially severe metabolic risk of hypoxia in seven of the 11 HIV-ART-exposed neonates included in this study (Moutloatse *et al.* 2016).

This NMR metabolomics study recognized the outstanding benefits of ART for HIV-positive pregnant women and their offspring as recommended by the WHO. Focussing also on the consequences of HIV-ART exposure to *individual* cases included in the cohort of cases studied, supported the concern as raised in several investigations that have indicated dysfunctional energy metabolism and mitochondrial DNA aberrations in some newborns exposed to fetal ART. The short- and long-term effects of HIV-ART exposure during pregnancy on infancy and childhood call for continuous basic and clinical investigations, substantiated by a review that summarized the indications that environmental processes influence the propensity to disease in adulthood operate during the peri-conceptual, fetal, and infant phases of life (Gluckman & Hanson 2004). Meticulous clinical monitoring is recommended for newborns whose HIV-positive mothers may have been exposed to ART during pregnancy. In addition, as these children increasingly form a substantial number of individuals at risk in society, there is a pressing need for longitudinal studies, on subjects from early infancy to childhood, following exposure. Birth is the preferred point of departure for longitudinal investigations on the consequences for the infants, children and even adults following the exposure of the fetus to the prenatal HIV-ART environment.

7.3.5 Future prospects

The main research lines followed by the six major NMR metabolomics studies on cord blood which were focused on in this review (Tea *et al.* 2012; Ivorra *et al.* 2012; Moutloatse *et al.* 2016; Reinke *et al.* 2013; Sanz-Cortés *et al.* 2013; McCance 2011), were: sample collection (cord blood/venous/arterial), preparation of samples and measurement of NMR spectra (500 and 600 MHz spectrometers), application of multivariate statistical techniques (mostly unsupervised PCA and unsupervised PLS-DA) to identify differences in spectral profiles in samples from “control” and “affected” newborns, limited metabolite identification and mostly expressed as semi-quantitative values (e.g. areas from normalized peak signals) and provisional rationalisation of the observed important/diagnostic metabolites related to some molecular pathways. These approaches are typical for exploratory inductive method used in novel applications of metabolomics technology (Goodacre *et al.* 2004). Exploratory research has proved to be valuable for researchers that have initiated programs on metabolomics in neonatology.

To a limited extent, this approach might be all that is required to answer some guiding questions of interest. However, more novel biochemical information, in-depth insights and critical perspectives on clinical questions can be realized through metabolomics investigations – and that is the aspiration expressed by neonatologists (Fanos *et al.* 2012, 2013, 2014; Mussap *et al.* 2013).

So, can NMR metabolomics studies contribute to the advancement of objective clinical information for diagnosis, prognosis and treatment monitoring in neonatology? All investigations reviewed here unequivocally indicated that the pathophysiological conditions investigated were reflected in differential metabolic profiles in cord blood and demonstrated that metabolic alterations are present and detectable at birth, providing metabolic information on the impact of the intrauterine environment on the life of the fetus. Hence, we formulate three perspectives, which we regard as key, as steps to take NMR metabolomics forward in the field of neonatology.

NMR has been proven to be amenable to analysis of an array of biological fluids.

Metabolite concentrations in biofluids from affected conditions reflect perturbations of the metabolic dynamic of the biological system due to innate genetic modification, pathophysiological stimuli or environmental factors. To characterize the multitude of metabolites found in any given biosample, NMR is but one technology among the wide

range of analytical platforms available for metabolomics studies. Each technology has advantages and disadvantages, and the choice for a given analytical platform mostly depends on the focus of the study and the samples (Emwas *et al.* 2015). The use of NMR in metabolomics is particularly attractive by providing a non-destructive and quantitative mode of analysis with high specificity and repeatability, requires minimal sample preparation and through its ability to identify novel compounds via unique one- or two-dimensional spectral patterns. NMR results also have a proven track record of translating *in vitro* findings into *in vivo* clinical applications. (Markley *et al.* 2017).

These virtues combine well with those of cord blood: it may be collected in a non-invasive way and used for rapid investigation of the biochemical status of a newborn at delivery, which may be linked to the maternal status and providing guidance to the design of strategies for prenatal and postnatal interventions. In terms of analytical limitations, the use of identical NMR tubes is required and within identical instrument and analytical parameters, factors such as temperature, chemical shift referencing, tuning, shimming, compensation for magnetic field drift and water suppression methods are critical for consistent and reproducible results (Bharti & Roy 2012). These conditions must be strictly adhered to from sample to sample to allow for rigorous quality assurance of the analysts (Mason *et al.* 2014) and robust intra-laboratory and inter-laboratory comparison. Variations in the analytical precision can have dramatic effects on the spectroscopic data, leading to an inability to interpret information generated at different clinical settings.

The normal cord blood metabolome

David Wisart, leader of the HMP, recently remarked: “*Despite the growing number of publications focused on diagnostic metabolite biomarkers appearing in the scientific literature, a surprisingly small number of these reports provide information on quantified metabolite levels.*” (Emwas *et al.* 2016). Accurate identification and quantification of normal metabolites which constitute a biofluid – by most scientists considered as a challenging and not an inspirational task – is essential for advancing biomarker discoveries to clinical practice. Nevertheless, it has now become apparent from many metabolomics studies that to fully interpret metabolomics data, significant knowledge of the normal metabolite profile of the biofluid studied is required. In particular, the range of normal/baseline metabolic profile, and how it may vary as a function of health and disease parameters, is currently only poorly understood for many biofluids, and especially so for cord blood as indicated in Table 7-1. Hence, we propose that metabolomics will deliver robust and interpretable insights only if, in a dedicated

manner, time, energy and resources are focused on identifying the baseline quantitative characteristics of the normal cord blood metabolome.

The transdisciplinary imperative

In contrast to traditional science, metabolomics is distinguished by being an inductive science, has a societal directedness and requires a culture of transdisciplinary cooperation (Madsena *et al.* 2010). Biology is the point of departure in metabolomics studies and new scholarly questions, not previously definable in the traditional sciences, can now be attempted. The mode of analyses to address these questions is, however, chemical and requires specialized technologies for measurement of the multitude of small molecules. Statistical methods required for analysis of such complex data likewise called for the development of new statistical, bioinformatics and data mining approaches, mostly of a multivariate nature. Interpretation of the functional meaning of the statistical information is beyond the ability of the statistics experts, but depends on biological scholars. Most importantly, the required transdisciplinary approach to provide the metabolomics outcome is sterile if pursued by experts in isolation: constant interaction and reflection between all experts is the absolute requirement

A key feature of transdisciplinarity is the practical unification of the meanings and insights that transpire from different disciplines. In metabolomics, the transdisciplinary approach is resolutely open in so far as it goes beyond the exact sciences and disciplines traditionally associated with the numerical (statistics; bioinformatics), physical (chemistry; chemometrics) and biological/clinical (biochemistry; neonatology) fields. Transdisciplinarity thus complements disciplinary approaches and does not strive for mastery of several disciplines but aims to open all disciplines to that which they share and to that which lies beyond them. Individual researchers should *“demonstrate disciplinary competence and know how concepts and methodologies from other disciplines relate to their own, having mastered some concepts therein. They are expected to be able to constructively communicate with those from other disciplines in a problem-focused manner”* (Stein *et al.* 2007). Transformation of the minds of the participants from all disciplines towards the characteristics and realities of metabolomics will be key towards its progress in general, and especially so in all fields of clinical applications with its progressive focus on personalized medicine (Collins & Varmus 2015). A recent quote from Fanos 2016 (Fanos *et al.* 2016) is worthwhile as a final note for personal reflection on the future of metabolomics: *“As for the future, your task is not to foresee, but to enable it”* (De Saint-Exupéry *et al.* 1950).

REFERENCES

SECTION 7.1 – 7.2:

Atherton, P. J., Smith, K., Etheridge, T., Rankin, D., & Rennie, M. J. (2010). Distinct anabolic signalling responses to amino acids in C2C12 skeletal muscle cells. *Amino Acids*, 38(5), 1533–1539.

Bell, J. D., Brown, J. C., Sadler, P. J., Garvie, D., Macleod, A. F., & Lowy, C. (1989). Maternal and cord blood plasma. Comparative analyses by ¹H NMR spectroscopy. *NMR in Biomedicine*, 2(2), 61–65.

Buonocore, G., Perrone, S., Longini, M., Vezzosi, P., Marzocchi, B., Paffetti, P., & Bracci, R. (2002). Oxidative stress in preterm neonates at birth and on the seventh day of life. *Pediatric Research*, 52(1), 46–49.

Croze, M. L., & Soulage, C. O. (2013). Potential role and therapeutic interests of myo-inositol in metabolic diseases. *Biochimie*, 95(10), 1811–1827.

Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., & Everett, J. R. (2016). A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Computational and Structural Biotechnology Journal*, 14, 135–153.

Eley, H. L., Russell, S. T., Baxter, J. H., Mukerji, P., & Tisdale, M. J. (2007). Signaling pathways initiated by β -hydroxy- β -methylbutyrate to attenuate the depression of protein synthesis in skeletal muscle in response to cachectic stimuli. *American Journal of Physiology-Endocrinology and Metabolism*, 293(4), E923–E931.

Fan, T. W. M., & Lane, A. N. (2016). Applications of NMR spectroscopy to systems biochemistry. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 92, 18–53.

Hashimoto, F., Nishiumi, S., Miyake, O., Takeichi, H., Chitose, M., Ohtsubo, H., et al. (2013). Metabolomics analysis of umbilical cord blood clarifies changes in saccharides associated with delivery method. *Early Human Development*, 89(5), 315–320.

Ivorra, C., García-Vicent, C., Chaves, F. J., Monleón, D., Morales, J. M., & Lurbe, E. (2012). Metabolomic profiling in blood from umbilical cords of low birth weight newborns. *Journal of Translational Medicine*, 10(1), 1–10.

Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43, 34–40.

Newell, M. L., & Bunders, M. J. (2013). Safety of antiretroviral drugs in pregnancy and breastfeeding for mother and child. *Current Opinion in HIV and AIDS*, 8(5), 504–510.

Pérez, G. M., Cox, V., Ellman, T., Moore, A., Patten, G., Shroufi, A., et al. (2016). 'I Know that I Do Have HIV but Nobody Saw Me': Oral HIV Self-Testing in an Informal Settlement in South Africa. *PloS One*, 11(4), 1–15.

Tea, I., Le Gall, G., Küster, A., Guignard, N., Alexandre–Gouabau, M. C., Darmaun, D., & Robins, R. J. (2012). ¹ H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*, 7(1), 1–12.

Williams, A., Jahreiss, L., Sarkar, S., Saiki, S., Menzies, F. M., Ravikumar, B., & Rubinsztein, D. C. (2006). Aggregate-prone proteins are cleared from the cytosol by autophagy: therapeutic implications. *Current Topics in Developmental Biology*, 76, 89–101.

WHO. (2015). Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. September 2015.

SECTION 7.3:

Alexandre-Gouabau, M. C., Courant, F., Moyon, T., Küster, A., Le Gall, G., Tea, I. et al. (2013). Maternal and cord blood LC-HRMS metabolomics reveal alterations in energy and polyamine metabolism, and oxidative stress in very-low birth weight infants. *Journal of Proteome Research*, 12(6), 2764–2778.

Barker, D. (2004). Developmental origins of adult health and disease. *Journal of Epidemiology and Community Health*, 58(2), 114.

Barker, D. J., Osmond, C., Forsén, T. J., Kajantie, E., & Eriksson, J. G. (2005). Trajectories of growth among children who have coronary events as adults. *New England Journal of Medicine*, 353(17), 1802–1809.

- Barret, B., Tardieu, M., Rustin, P., Lacroix, C., Chabrol, B., Desguerre, I., et al. (2003). Persistent mitochondrial dysfunction in HIV-1-exposed but uninfected infants: clinical screening in a large prospective cohort. *AIDS*, 17(12), 1769–1785.
- Battaglia, F. C., & Meschia, G. (1986). An introduction to fetal physiology. *Academic Press*.
- Bell, J. D., Brown, J. C., Sadler, P. J., Garvie, D., Macleod, A. F., & Lowy, C. (1989). Maternal and cord blood plasma. Comparative analyses by ¹H NMR spectroscopy. *NMR in Biomedicine*, 2(2), 61–65.
- Bernstein, I. M., Horbar, J. D., Badger, G. J., Ohlsson, A., Golan, A., & Vermont Oxford Network. (2000). Morbidity and mortality among very-low-birth-weight neonates with intrauterine growth restriction. *American Journal of Obstetrics and Gynecology*, 182(1), 198–206.
- Bertram, H. C., Eggers, N., & Eller, N. (2009). Potential of human saliva for nuclear magnetic resonance-based metabolomics and for health-related biomarker identification. *Analytical Chemistry*, 81(21), 9188–9193.
- Bharti, S. K.; & Roy, R. (2012). Quantitative H-1 NMR spectroscopy. Trends in Analytical Chemistry, 35, 5–26.
- Blanche, S., Tardieu, M., Rustin, P., Slama, A., Barret, B., Firtion, G., et al. (1999). Persistent mitochondrial dysfunction and perinatal exposure to antiretroviral nucleoside analogues. *Lancet*, 354, 1084–1089.
- Bouatra, S., Aziat, F., Mandal, R., Guo, A. C., Wilson, M. R., Knox, C., et al. (2013). The human urine metabolome. *PloS One*, 8(9), e73076.
- Brogly, S. B., Ylitalo, N., Mofenson, L. M., Oleske, J., Van Dyke, R., Crain, M. J., et al. (2007). In utero nucleoside reverse transcriptase inhibitor exposure and signs of possible mitochondrial dysfunction in HIV-uninfected children. *AIDS*, 21(8), 929–938.
- Burdge, G. C., & Lillycrop, K. A. (2010). Nutrition, epigenetics, and developmental plasticity: implications for understanding human disease. *Annual Review of Nutrition*, 30, 315–339.

Chasela, C. S., Hudgens, M. G., Jamieson, D. J., Kayira, D., Hosseinipour, M. C., Kourtis, A. P., et al. (2010). Maternal or infant antiretroviral drugs to reduce HIV-1 transmission. *The New England Journal of Medicine*, 362(24), 2271–2281.

Collins, F. S. & Varmus, H. (2015). A New Initiative on Precision Medicine. *New England Journal of Medicine*, 372, 793–795.

Committee on Practice Bulletins--Gynecology, A. (2001). Intrauterine growth restriction. Clinical management guidelines for obstetrician-gynecologists. American College of Obstetricians and Gynecologists. *International Journal of Gynaecology and Obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*, 72(1), 85.

Courant, F., Antignac, J. P., Monteau, F., & Le Bizec, B. (2013). Metabolomics as a potential new approach for investigating human reproductive disorders. *Journal of Proteome Research*, 12(6), 2914–2920.

Dani, C., Bresci, C., Berti, E., Ottanelli, S., Mello, G., Mecacci, F., et al. (2014). Metabolomic profile of term infants of gestational diabetic mothers. *The Journal of Maternal-Fetal & Neonatal Medicine*, 27(6), 537–542.

Davison, J. M., Mellott, T. J., Kovacheva, V. P., & Blusztajn, J. K. (2009). Gestational choline supply regulates methylation of histone H3, expression of histone methyltransferases G9a (Kmt1c) and Suv39h1 (Kmt1a), and DNA methylation of their genes in rat fetal liver and brain. *Journal of Biological Chemistry*, 284(4), 1982–1989.

De Saint-Exupéry, A. (1950). *The wisdom of the sands*. New York: Hartcourt, Brace.

Dessi, A., Atzori, L., Noto, A., Visser, G. H. A., Gazzolo, D., Zanardo, V. et al. (2011). Metabolomics in newborns with intrauterine growth retardation (IUGR): urine reveals markers of metabolomics syndrome. *The Journal of Maternal-Fetal and Neonatal Medicine*, 24(Supplementary 2), 35–39.

Dessi, A., Marincola, F. C., & Fanos, V. (2015). Metabolomics and the great obstetrical syndromes—GDM, PET, and IUGR. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 29(2), 156–164.

Diaz, S. O., Barros, A. S., Goodfellow, B. J., Duarte, I. F., Galhano, E., Pita, C., et al. (2013). Second trimester maternal urine for the diagnosis of trisomy 21 and prediction of poor pregnancy outcomes. *Journal of Proteome Research*, 12(6), 2946–2957.

Diaz, S. O., Pinto, J., Grac[□]a, G., Duarte, I. F., Barros, A. S., Galhano, E., et al. (2011). Metabolic biomarkers of prenatal disorders: an exploratory NMR metabonomics study of second trimester maternal urine and blood plasma. *Journal of Proteome Research*, 10(8), 3732–3742.

Diderholm, B. (2009). Perinatal energy metabolism with reference to IUGR & SGA: studies in pregnant women & newborn infants. *Indian Journal of Medical Research*, 130(5), 612–617.

Dunn, W. B., Brown, M., Worton, S. A., Davies, K., Jones, R. L., Kell, D. B., & Heazell, A. E. (2012). The metabolome of human placental tissue: investigation of first trimester tissue and changes related to preeclampsia in late pregnancy. *Metabolomics*, 8(4), 579–597.

Emwas, A. H. M. (2015). The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Metabonomics: Methods and Protocols*, 161–193.

Emwas, A. H., Roy, R., McKay, R. T., Ryan, D., Brennan, L., Tenori, L., et al. (2016). Recommendations and standardization of biomarker quantification using NMR-based metabolomics with particular focus on urinary analysis. *Journal of Proteome Research*, 15(2), 360–373.

Engelke, U. F. H., Kremer, B., Kluijtmans, L. A. J., van der Graaf, M., Morava, E., Loupatty, F. J., et al. (2006). NMR spectroscopic studies on the late onset form of 3-methylglutaconic aciduria type I and other defects in leucine metabolism. *NMR in Biomedicine*, 19, 271–278.

Engelke, U. F. H., Sass, J. O., Van Coster, R. N., Gerlo, E., Olbrich, H., Krywawych, S., et al. (2008). NMR spectroscopy of aminoacylase 1 deficiency, a novel inborn error of metabolism. *NMR in Biomedicine*, 21, 138–147.

Fanos, V. (2016). 10 P Pediatrics: Notes for the future. *Journal of Pediatric and Neonatal Individualized Medicine*, 5(1), 1–4.

Fanos, V., Antonucci, R., Barberini, L., Noto, A., & Atzori, L. (2012). Clinical application of metabolomics in neonatology. *The Journal of Maternal-Fetal & Neonatal Medicine*, 25(Supplementary 1), 104–109.

Fanos, V., Atzori, L., Makarenko, K., Melis, G. B., & Ferrazzi, E. (2013). Metabolomics application in maternal-fetal medicine. *BioMed Research International*, 2013, 1–9.

Fanos, V., Mussap, M., Faa, G., & Papageorgiou, A. (2014). The next ten years in neonatology: new directions in research. *Journal of Pediatric and Neonatal Individualized Medicine*, 3(2), 1–6.

Fattuoni, C., Palmas, F., Noto, A., Fanos, V., & Barberini, L. (2015). Perinatal asphyxia: a review from a metabolomics perspective. *Molecules*, 20(4), 7000–7016.

Gluckman, P. D., & Hanson, M. A. (2004). Living with the past: evolution, development, and patterns of disease. *Science*, 305, 1733–1736.

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in utero and early-life conditions on adult health and disease. *New England Journal of Medicine*, 359(1), 61–73.

Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G., & Kell, D. B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in biotechnology*, 22(5), 245–252.

Graça, G., Duarte, I. F., Barros, A. S., Goodfellow, B. J., Diaz, S. O., Pinto, J., et al. (2010). Impact of prenatal disorders on the metabolic profile of second trimester amniotic fluid: a nuclear magnetic resonance metabonomic study. *Journal of Proteome Research*, 9(11), 6016–6024.

Graça, G., Duarte, I. F., Barros, A. S., Goodfellow, B. J., Diaz, S., Carreira, I. M., et al. (2009). ¹H NMR based metabonomics of human amniotic fluid for the metabolic characterization of fetus malformations. *Journal of Proteome Research*, 8(8), 4144–4150.

Graça, G., Duarte, I. F., Goodfellow, B. J., Barros, A. S., Carreira, I. M., Couceiro, A. B., et al. (2007). Potential of NMR spectroscopy for the study of human amniotic fluid. *Analytical Chemistry*, 79(21), 8367–8375.

Hay Jr, W. W. (2012). Care of the infant of the diabetic mother. *Current Diabetes Reports*, 12(1), 4–15.

Huynh, J., Xiong, G., & Bentley-Lewis, R. (2014). A systematic review of metabolite profiling in gestational diabetes mellitus. *Diabetologia*, 57(12), 2453–2464.

- Isganaitis, E., Jimenez-Chillaron, J., Woo, M., Chow, A., DeCoste, J., Vokes, M., et al. (2009). Accelerated postnatal growth increases lipogenic gene expression and adipocyte size in low-birth weight mice. *Diabetes*, 58(5), 1192–1200.
- Ivorra, C., García-Vicent, C., Chaves, F. J., Monleón, D., Morales, J. M., & Lurbe, E. (2012). Metabolomic profiling in blood from umbilical cords of low birth weight newborns. *Journal of Translational Medicine*, 10(1), 1–10.
- Jao, J., & Abrams, E. J. (2014). Metabolic complications of in utero maternal HIV and antiretroviral exposure in HIV-exposed infants. *The Pediatric Infectious Disease Journal*, 33(7), 734–740.
- Kaser, S., Ebenbichler, C. F., Wolf, H. J., Sandhofer, A., Stanzl, U., Ritsch, A., & Patsch, J. R. (2001). Lipoprotein profile and cholesteryl ester transfer protein in neonates. *Metabolism*, 50(6), 723–728.
- Katsila, T., & Patrinos, G. P. (2015). The implications of metabolotypes for rationalizing therapeutics in infants and children. *Frontiers in Pediatrics*, 3, 1–3.
- Lawn, J. E., Cousens, S., Zupan, J., & Lancet Neonatal Survival Steering Team. (2005). 4 million neonatal deaths: when? Where? Why? *The Lancet*, 365, 891–900.
- Liotta, G., Mancinelli, S., Nielsen-Saines, K., Gennaro, E., Scarcella, P., Magid, N. A., et al. (2013). Reduction of maternal mortality with highly active antiretroviral therapy in a large cohort of HIV-infected pregnant women in Malawi and Mozambique. *PLoS One*, 8(8), 1–7.
- Lowe, W. L., & Karban, J. (2014). Genetics, genomics and metabolomics: new insights into maternal metabolism during pregnancy. *Diabetic Medicine*, 31(3), 254–262.
- Madsena, R., Lundstedt, T. & Trygg, J. (2010). Chemometrics in metabolomics—A review in human disease diagnosis. *Analytica Chimica Acta*, 659, 23–33.
- Markley, J. L., Brüschweiler, R., Edison, A. S., Eghbalnia, H. R., Powers, R., Raftery, D., & Wishart, D. S. (2017). The future of NMR-based metabolomics. *Current Opinion in Biotechnology*, 43, 34–40.

Mason, S., P Moutloatse, G., Marceline van Furth, A., Solomons, R., van Reenen, M., Reinecke, C., & Koekemoer, G. (2014). KEMREP: A new qualitative method for the assessment of an analyst's ability to generate a metabolomics data matrix by gas chromatography–mass spectrometry. *Current Metabolomics*, 2(1), 15–26.

McCance, D. R. (2011). Pregnancy and diabetes. *Best Practice & Research Clinical Endocrinology & Metabolism*, 25(6), 945–958.

Mehedint, M. G., Niculescu, M. D., Craciunescu, C. N., & Zeisel, S. H. (2010). Choline deficiency alters global histone methylation and epigenetic marking at the Re1 site of the calbindin 1 gene. *The FASEB Journal*, 24(1), 184–195.

Montoya, D. A., White, A. M., Williams, C. L., Blusztajn, J. K., Meck, W. H., & Swartzwelder, H. S. (2000). Prenatal choline exposure alters hippocampal responsiveness to cholinergic stimulation in adulthood. *Developmental Brain Research*, 123(1), 25–32.

Moutloatse, G. P., Bunders, M. J., van Reenen, M., Mason, S., Kuijpers, T. W., Engelke, U. F. H. & Reinecke, C. J. (2016). Metabolic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood. *Metabolomics*, 12, 175–189.

Mussap, M., Antonucci, R., Noto, A., & Fanos, V. (2013). The role of metabolomics in neonatal and pediatric laboratory medicine. *Clinica Chimica Acta*, 426, 127–138.

Newell, M. L., & Bunders, M. J. (2013). Safety of antiretroviral drugs in pregnancy and breastfeeding for mother and child. *Current Opinion in HIV and AIDS*, 8(5), 504–510.

Newgard, C. B., An, J., Bain, J. R., Muehlbauer, M. J., Stevens, R. D., Lien, L. F., et al. (2009). A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metabolism*, 9(4), 311–326.

Novak, M., Monkus, E. F., Chung, D., & Buch, M. (1981). Carnitine in the perinatal metabolism of lipids I. Relationship between maternal and fetal plasma levels of carnitine and acylcarnitines. *Pediatrics*, 67(1), 95–100.

Pecks, U., Brieger, M., Schiessl, B., Bauerschlag, D. O., Piroth, D., Bruno, B., et al. & Rath, W. (2012). Maternal and fetal cord blood lipids in intrauterine growth restriction. *Journal of Perinatal Medicine*, 40(3), 287–296.

- Pinto, J., Almeida, L. M., Martins, A. S., Duarte, D., Barros, A. S., Galhano, E., et al. (2015). Prediction of gestational diabetes through NMR metabolomics of maternal blood. *Journal of Proteome Research*, 14(6), 2696–2706.
- Platt, M. W., & Deshpande, S. (2005, August). Metabolic adaptation at birth. *In Seminars in Fetal and Neonatal Medicine*, 10(4), 341–350.
- Prefumo, F., Gaze, D. C., Papageorgiou, A. T., Collinson, P. O., & Thilaganathan, B. (2007). First trimester maternal serum ischaemia-modified albumin: a marker of hypoxia-ischaemia-driven early trophoblast development. *Human Reproduction*, 22(7), 2029–2032.
- Psychogios, N., Hau, D. D., Peng, J., Guo, A. C., Mandal, R., Bouatra, S., et al. (2011). The Human Serum Metabolome. *PloS One*, 6(2), 1–23.
- Raff, H., Bruder, E. D., Jankowski, B. M., & Goodfriend, T. L. (2000). Neonatal hypoxic hyperlipidemia in the rat: effects on aldosterone and corticosterone synthesis in vitro. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 278(3), R663–R668.
- Reinke, S. N., Walsh, B. H., Boylan, G. B., Sykes, B. D., Kenny, L. C., Murray, D. M., & Broadhurst, D. I. (2013). ¹H NMR derived metabolomic profile of neonatal asphyxia in umbilical cord serum: implications for hypoxic ischemic encephalopathy. *Journal of Proteome Research*, 12(9), 4230–4239.
- Rodiño-Janeiro, B. K., González-Peteiro, M., Ucieda-Somoza, R., González-Juanatey, J. R., & Álvarez, E. (2010). Glycated albumin, a precursor of advanced glycation end-products, up-regulates NADPH oxidase and enhances oxidative stress in human endothelial cells: molecular correlate of diabetic vasculopathy. *Diabetes/Metabolism Research and Reviews*, 26(7), 550–558.
- Sachse, D., Sletner, L., Mørkrid, K., Jenum, A. K., Birkeland, K. I., Rise, F., ... & Berg, J. P. (2012). Metabolic changes in urine during and after pregnancy in a large, multiethnic population-based cohort study of gestational diabetes. *PloS One*, 7(12), 1–12.

- Sanz-Cortés, M., Carbajo, R. J., Crispi, F., Figueras, F., Pineda-Lucena, A., & Gratacós, E. (2013). Metabolomic profile of umbilical cord blood plasma from early and late intrauterine growth restricted (IUGR) neonates with and without signs of brain vasodilation. *PloS One*, 8(12), 1–21.
- Saugstad, O. D. (1975). Hypoxanthine as a measurement of hypoxia. *Pediatric Research*, 9(4), 158–161.
- Say, L., Chou, D., Gemmill, A., Tunçalp, Ö., Moller, A. B., Daniels, J., et al. (2014). Global causes of maternal death: a WHO systematic analysis. *The Lancet Global Health*, 2(6), e323–e333.
- Stein, Z. (2007). Modeling the Demands of Interdisciplinarity: Toward a Framework for Evaluating Interdisciplinary *Endeavors Integral Review*, 4(1), 91–107.
- Tea, I., Le Gall, G., Küster, A., Guignard, N., Alexandre–Gouabau, M. C., Darmaun, D., & Robins, R. J. (2012). 1 H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*, 7(1), 1–12.
- Walsh, B. H., Broadhurst, D. I., Mandal, R., Wishart, D. S., Boylan, G. B., Kenny, L. C., & Murray, D. M. (2012). The metabolomic profile of umbilical cord blood in neonatal hypoxic ischaemic encephalopathy. *PLoS One*, 7(12), 1–12.
- Wang, T. J., Larson, M. G., Vasan, R. S., Cheng, S., Rhee, E. P., McCabe, E., et al. (2011). Metabolite profiles and the risk of developing diabetes. *Nature Medicine*, 17(4), 448–453.
- Wevers, R. A., Engelke, U., Wendel, U., de Jong, J. G. N., Gabreeëls, F. J. M., & Heerschap, A. (1995). Standardized method for highresolution 1H-NMR of cerebrospinal fluid. *Clinical Chemistry*, 41(5), 744–751.
- Wishart, D. S., Lewis, M. J., Morrissey, J. A., Flegel, M. D., Jeroncic, K., Xiong, Y., et al. (2008). The human cerebrospinal fluid metabolome. *Journal of Chromatography B*, 871(2), 164–173.
- Wishart, D. S., Tzur, D., Knox, C., Eisner, R., Guo, A. C., Young, N., et al. (2007). HMDB: the human metabolome database. *Nucleic Acids Research*, 35(Supplementary 1), D521–D526.

ANNEXURE 1:

Supplementary information:

Metabolic risks at birth of neonates exposed *in utero* to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood

A1.1 Sample preparation and spectral data

Samples were prepared and spectra generated according to the workflow shown in Figure 4-1 (stage 1) and by using the standard operational procedure established by the Nijmegen Translational Metabolic Laboratory (Engelke *et al.* 2006, 2008; Wevers *et al.* 1995). Plasma and serum samples contain large molecular weight proteins and lipoproteins, which affect the NMR spectroscopic identification of smaller molecular weight metabolites. This interference can be avoided by deproteinization using ultra-filtration (Sartorius Centrisart 1 10,000 MWCO Sartorius AG, Göttingen, Germany — molecular mass cutoff at 10 kDa). Prior to filtration, the filters were washed twice with 6 mL of distilled water and centrifuged (3000 x g, 10 min) to remove the glycerol bound to the filter membrane. A sample volume of 1 mL (minimum of 0.8 mL) was aliquoted into the cleaned filters and centrifuged (3000 x g, 15 min). Once deproteinized, 700 μ L of the filtered supernatant was transferred into a test tube and diluted with 20 μ L of internal standard solution [freshly prepared deuterium oxide (D_2O) solution containing 20.2 mmol/L 3-trimethylsilyl-tetradeuteriopropionic acid (TSP, sodium salt; Merck, Darmstadt, Germany)] as the chemical shift reference ($\delta = 0.00$ ppm) and as a deuterium lock signal. Sample filtrates less than 700 μ L were diluted with an appropriate amount of distilled water to obtain a final sample volume of 700 μ L — the dilution factor was accounted for in calculating the concentration of metabolites. The pH of the samples was then altered to 2.50 ± 0.05 using HCl and NaOH at room temperature. A final volume of 650 μ L of the prepared samples was placed in 5-mm high quality NMR tubes (Wilmad Royal Imperial; Aldrich) for analysis.

High-resolution NMR spectra of all samples (Fig. A1-1) were recorded using a 500 MHz ^1H resonance frequency spectrometer (AMX-500; Bruker Analytische Messtechnik, Karlsruhe, Germany) with a triple-resonance inverse (TXI) $^1H \{^{15}N, ^{13}C\}$ probe-head equipped with X, Y, Z gradient coils and a sample changer. The samples were loaded in a random order to avoid any confounding machine variation. All spectra were recorded at a temperature of 298 K. NMR data pre-processing was performed

using Bruker Topspin (V3.1) and Bruker AMIX (V3.9.12) for binning and quantification (Figure 4-1: stage 2). Phase correction was performed manually by adjusting the baseline peaks according to the TSP calibrated at $\delta = 0.00$ ppm. Integration into bins (or buckets) across the spectral regions was performed automatically in AMIX through variable-sized bucketing (van den Berg *et al.* 2006; Eriksson *et al.* 2006; Cui & Churchill 2003; Best *et al.* 2005; De Meyer *et al.* 2008). Spectral regions devoid of peaks (noise) and the region containing water resonances were excluded from the data. All data were normalized in AMIX by dividing each integrated segment by the total spectrum intensity (minus the excluded region). Normalization, i.e. making the total integrated intensity of each spectrum the same, compensated for the differences in overall concentration between individual plasma samples.

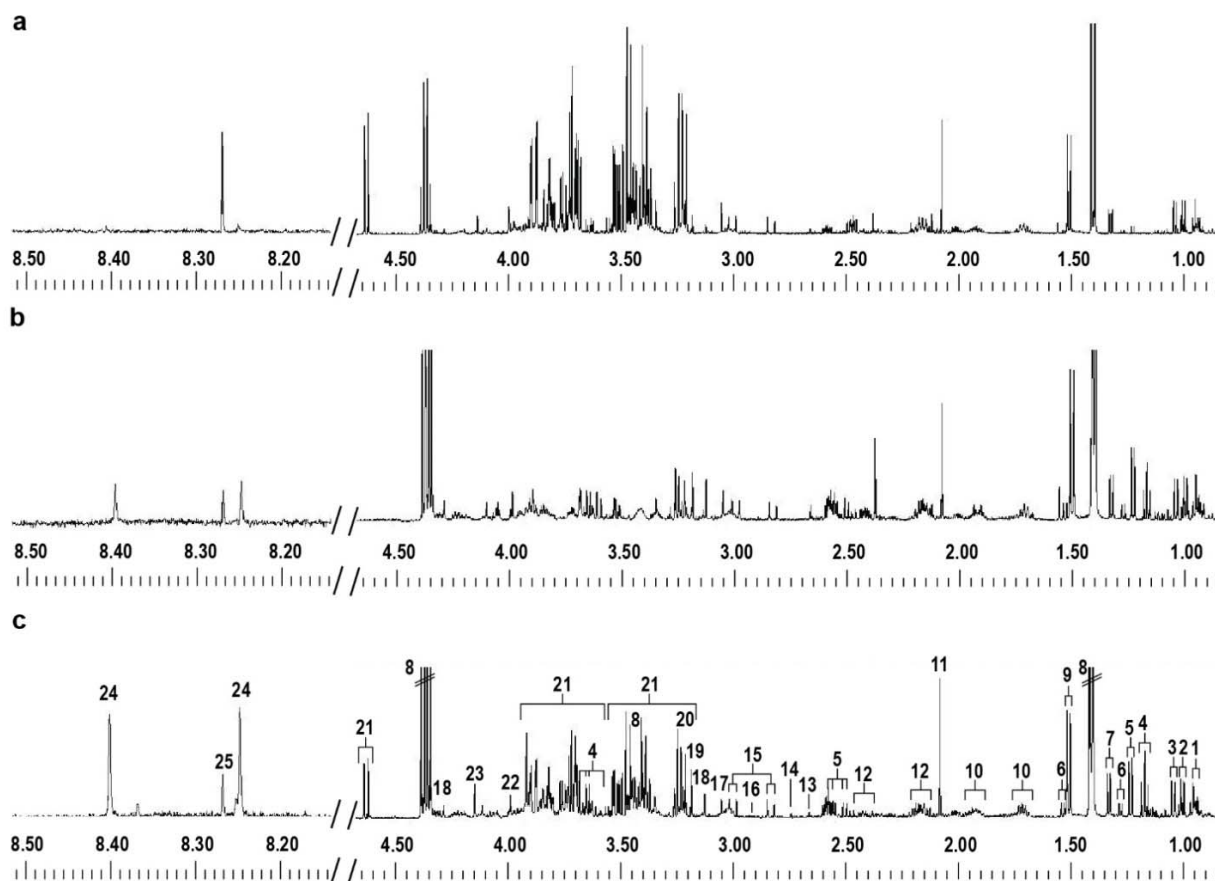


Figure A1-1: 500 MHz ^1H -NMR spectra from one representative sample of infant serum (a), cord blood plasma from a newborn unexposed to HIV-ART *in utero* (b) and from a neonate born to a mother who received HIV-ART (c). Samples were analysed at pH 2.5 ± 0.05 and the internal standard [trimethylsilyl-propionic acid (TSP)] was positioned at 0.00 ppm. Assignments (ppm) of important metabolites and a few unknowns identified in the samples from the HIV-ART-exposed neonates are shown in (c): 1, isoleucine/leucine (0.94t/0.95d); 2, valine/isoleucine (1.00d/1.01d); 3, valine (1.04d); 4, ethanol (1.17t, 3.64q); 5, 3-hydroxybutyric acid (1.23d, 2.53AB); 6, ^{13}C lactic acid (1.30d, 1.52d); 7, threonine (1.33d); 8, lactic acid (1.41d, 4.36q); 9, alanine (1.51d); 10, lysine (1.73m, 1.92m); 11, acetic acid (2.08s); 12, glutamine (2.16m, 2.47m); 13, succinic acid (2.66s); 14, sarcosine (2.74s); 15, citric acid (2.91AB); 16, dimethylglycine (2.93s); 17, creatine (3.05s); 18, creatinine (3.13s, 4.29s); 19, acetylcarnitine/choline (3.19s); 20, carnitine (3.22s); 21, glucose (3.23dd, 3.3-3.9, 4.64d, 5.22d); 22, unknown (3.99s); 23, unknown (4.15s); 24, hypoxanthine (8.21s, 8.42s); 25, formic acid (8.24s). The region containing phenylalanine (7.32m) and tyrosine (6.89dd, 7.19dd) is not included in the figure. Multiple peaks are designated by: s singlet, d doublet, dd double doublet, t triplet, m multiplet, AB AB-system.

The normalized data for each bin were amalgamated into one ($n \times p$) data matrix composed of p bins (indicated by their ppm positions and variable name — shown as unknown or assigned name) and n quantities as spectral amplitude or area in Excel format (X), shown for each individual experimental case. The data matrix follows the format as shown below and was then used for statistical pattern recognition and to generate metabolomics information:

$$X_{(n \times p)} = \begin{bmatrix} X_{1-1} & X_{1-2} & \dots & X_{1-p} \\ X_{2-1} & X_{2-2} & \dots & X_{2-p} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n-1} & X_{n-2} & \dots & X_{n-p} \end{bmatrix}$$

The original data for the 114 bins are shown for the three experimental groups in Tables A1-2 (HIV-ART exposed newborns), A1-3 (unexposed newborns) and A1-4 (infants). The resulting data matrix (peak integral values/bins per sample) was used for analysis by pattern recognition methods, followed by bin reduction to provide a matrix consisting of only variables which were unequivocally assigned to metabolites. Table A1-1 summarizes the clinical parameters for the mothers of the ART-exposed neonates (Table A1-2) along with details of their clinical follow-up during the pregnancy. Their CD4+ T-cell count and plasma HIV RNA were measured during the first, second and third trimester of pregnancy, along with records of the type of cART regimen the woman received including, their start and stop dates.

Table A1–1: Clinical information of the HIV-positive mothers receiving ART during pregnancy.

Age (years)	Ethnicity	Viral load ^a			CD4 count (10E ⁶ cells/L)			ART treatment regime (Y: Yes; N: No)						
		1 st	2nd	3rd	Nadir ^b	1st	2nd	3rd	Zidovudine + Lamivudine	Lopinavi r + Ritonavir	Lamivudine	Nelfinavi r	Nevirapine	Tenofovir
39	Caucasian	0	0	0	180	420	380	380	N	Y	N	N	Y	N
34	Black	0	0	0	200	540	460	500	Y	N	N	Y	N	N
21	Black	0	0	0	130	290	260	280	Y	N	N	N	Y	N
24	Black	0	0	0	160	600	680	580	Y	Y	N	Y	N	N
39	Black	0	0	0	330	490	510	660	N	N	Y	N	Y	N
24	Black	0	0	523	800	n/a	800	750	Y	Y	N	N	N	N
40	Black	0	0	0	n/a	790	720	850	N	N	Y	N	Y	N
29	Black	0	0	0	180	610	430	550	Y	N	Y	N	Y	Y
39	Black	28365	16210	0	180	180	230	200	Y	Y	N	N	N	N
30	Black	0	0	0	40	460	400	510	Y	N	N	N	Y	N
33	Black	0	0	0	170	440	430	410	Y	Y	N	N	N	Y
34	Caucasian	0	0	0	220	1300	970	950	Y	N	N	N	Y	N

^a Viral load: 0 = < 50 copies/mL

^b Nadir is lowest CD4 count

n/a, not applicable

Table A1–2: ¹H-NMR spectral data for 114 bins relating to cord blood from 12 neonates exposed to HIV-ART.

Assignments: ppm and corresponding metabolite names	HIV1	HIV2	HIV3	HIV4	HIV5	HIV6	HIV7	HIV8	HIV9	HIV10	HIV11	HIV12
0.846	3.987E-05	9.3E-05	0.00011	0.0001	0.000197	-4.1E-07	-1.63E-06	0.00017	0.000217	0.000267	0.000163	0.000135
0.861	3.251E-05	3.6E-05	2.9E-05	3E-05	3.8E-05	0.0000247	1.502E-05	7.4E-05	7.5E-05	7.5E-05	4.75E-05	7.45E-05
0.865	0.0001537	6.6E-05	0.00013	5E-05	7.68E-05	9.491E-05	3.155E-05	0.0001	9.83E-05	7.58E-05	3.54E-05	0.000174
0.869 (2-OH Isovaleric acid)	0.0001304	8.5E-05	0.00014	0.0002	0.000263	8.879E-05	5.227E-05	0.00016	0.000111	0.000171	0.00011	0.000159
0.889	0.0015895	0.00115	0.00184	0.0019	0.001643	0.0019608	0.0003131	0.00169	0.001563	0.001497	0.000989	0.002472
0.917 (Ethylmalonic acid)	0.0014532	0.00086	0.00141	0.0008	0.000803	0.0013801	0.0004428	0.00127	0.001132	0.000958	0.000551	0.001786
0.925	0.0004583	0.00055	0.00042	0.0004	0.000404	0.0006063	0.0003174	0.00052	0.00059	0.000542	0.000479	0.000402
0.939 (Isoleucine/Leucine)	0.0060583	0.00524	0.00602	0.0059	0.007543	0.0065191	0.0033542	0.00612	0.005986	0.00603	0.005169	0.006172
0.960 (Leucine)	0.0061628	0.00657	0.00629	0.0055	0.007348	0.0068773	0.0036791	0.00671	0.006879	0.005831	0.007022	0.005424
1.004 (Isoleucine/Valine)	0.0100673	0.01071	0.01392	0.0092	0.014614	0.0123229	0.0061602	0.01062	0.011031	0.009778	0.008366	0.010442
1.014 (Valine)	0.0068227	0.00731	0.00968	0.0064	0.010556	0.0083072	0.0043742	0.00757	0.007789	0.006913	0.005586	0.007408
1.065	0.0002081	0.00024	0.00036	0.0002	0.000291	0.0002681	0.0000714	0.00028	0.000201	0.000208	0.000194	0.000295
1.080	0.0008003	0.0006	0.00104	0.0006	0.000608	0.0008431	0.000234	0.00081	0.000419	0.000516	0.000415	0.001129
1.094	0.0005763	0.0004	0.00072	0.0003	0.00029	0.0005556	0.0001637	0.00054	0.000286	0.000378	0.000257	0.000862
1.104	0.0003278	0.0002	0.00049	2E-05	2.22E-05	0.000373	-7.66E-06	0.00024	0.000283	0.000143	3.57E-05	0.000703
1.113	3.881E-05	3E-05	0.00014	0.0002	0.000279	0.0000665	0.000056	0.00011	7.52E-05	0.000164	5.72E-05	0.000159
1.122	0.0008357	0.00103	0.00154	0.0009	0.001045	0.0013241	0.0002959	0.00093	0.000937	0.000864	0.000666	0.001579
1.130 (Propylene glycol)	0.0005378	0.00012	0.00038	0.0006	0.000754	0.0002236	0.000294	0.00041	0.000113	0.000415	0.00025	0.000654
1.136	0.0004322	0.00077	0.00098	0.0008	0.000814	0.0008633	0.0002522	0.00057	0.000603	0.000564	0.000549	0.000752
1.144 (Propylene glycol)	0.0005031	0.00011	0.00039	0.0005	0.000631	0.0002261	0.0002368	0.00043	8.94E-05	0.000404	0.000227	0.000649
1.170 (Ethanol)	0.0065846	0.006	0.00997	0.0106	0.013171	0.00968	0.0056679	0.00948	0.008082	0.004041	0.007337	0.010785
1.230 (3-OH-Butyric acid)	0.0016856	0.00892	0.00132	0.0095	0.001274	0.0044756	0.0030901	0.01389	0.003073	0.013309	0.001868	0.001379
1.275 (C13 Lactic acid)	0.0026506	0.00234	0.00355	0.0026	0.002885	0.0026047	0.0010339	0.00262	0.002453	0.001918	0.002011	0.002959
1.310	6.111E-05	0.0005	0.00049	0.0004	0.000341	0.0002691	5.003E-05	0.00022	0.000317	0.000346	0.000211	0.000346
1.321 (Threonine/3-hydroxyisovaleric acid)	0.0040476	0.0051	0.00558	0.0041	0.004219	0.0064289	0.0024974	0.00374	0.004472	0.005027	0.002139	0.00519

1.334 (Threonine)	0.0039119	0.00472	0.00536	0.0039	0.003966	0.0061944	0.0024241	0.00365	0.004299	0.004863	0.00188	0.005029
1.357	-7.97E-05	0.00043	0.0007	0.0003	0.000286	0.0002166	1.498E-05	0.00019	0.000257	0.000249	0.000254	0.000261
1.408 (Lactic acid)	0.4731585	0.29895	0.47249	0.3783	0.409481	0.4006067	0.2011428	0.39745	0.358212	0.244883	0.315172	0.446162
1.468	0.0018461	0.00076	0.00092	0.0006	0.000767	0.0010332	0.0002423	0.00087	0.000819	0.000889	0.000786	0.001266
1.507 (Alanine)	0.0286485	0.02871	0.03086	0.02	0.023026	0.0252337	0.0108369	0.02064	0.021306	0.017067	0.015155	0.033925
1.536 (C13 Lactic)	0.003778	0.00227	0.00284	0.0026	0.002715	0.0031533	0.0011309	0.00283	0.002641	0.001753	0.002269	0.003077
1.561 (Pyruvic acid)	0.0014314	0.00087	0.00228	0.0001	9.28E-05	0.0013763	5.662E-05	0.00052	0.000765	0.000176	0.000274	0.002219
1.722 (Lysine)	0.0154957	0.01442	0.01501	0.0115	0.013861	0.016866	0.0043725	0.01348	0.015243	0.014419	0.009605	0.015404
1.916 (Lysine/Glutamic acid)	0.0161892	0.01391	0.01638	0.0109	0.013627	0.0166767	0.0036403	0.0127	0.014604	0.014091	0.007914	0.018378
2.024 (Proline)	0.0070039	0.00734	0.00752	0.0056	0.006385	0.0076988	0.0015987	0.00593	0.00703	0.006189	0.005031	0.008859
2.080 (Acetic acid)	0.0197438	0.00407	0.01477	0.0101	0.007842	0.0105622	0.0066757	0.01098	0.002862	0.009482	0.012488	0.021942
2.124 (Methionine/Acetylcarnitine)	0.0017037	0.0016	0.00166	0.0008	0.000776	0.001395	0.0003555	0.00116	0.001509	0.000921	0.000608	0.001596
2.171 (Glutamine)	0.0174237	0.02163	0.01484	0.0118	0.013589	0.0210251	0.0053433	0.01145	0.016847	0.013585	0.009873	0.021381
2.219 (Acetone)	0.0011615	0.00301	0.00055	0.0005	0.000569	0.0007681	0.0004104	0.00049	0.001228	0.003186	0.000269	0.000448
2.332 (Proline/Unknown)	0.0033222	0.00374	0.00369	0.0029	0.003936	0.0035204	0.0009066	0.00303	0.003233	0.003377	0.002487	0.004434
2.378 (Pyruvic acid)	0.0024694	0.00177	0.00502	0.0006	0.000593	0.0027162	0.0003052	0.00123	0.001702	0.000676	0.000477	0.005033
2.396	0.0002401	0.00024	0.00031	0.0002	0.000311	0.0003112	0.0001454	0.00025	0.000232	0.000274	0.00013	0.000451
2.402	0.0005907	0.00038	0.00062	0.0004	0.000433	0.0006865	0.0001269	0.00044	0.00019	0.00033	0.000248	0.00129
2.410	0.0004707	0.00019	0.00039	0.0004	0.000386	0.0005493	0.0001137	0.00033	0.000231	0.000365	0.00014	0.001228
2.415 (Glutaric acid/Unknown)	0.0004143	0.00024	0.00046	0.0003	0.000323	0.0005115	4.508E-05	0.00035	0.000189	0.000238	0.000208	0.000921
2.423	0.0010597	0.00068	0.0011	0.0009	0.000778	0.001388	0.0001992	0.00082	0.000567	0.00078	0.000408	0.002925
2.429 (Glutaric acid)	0.0004129	0.00026	0.00046	0.0004	0.000346	0.0005532	0.0001023	0.00034	0.000231	0.000255	0.000207	0.000967
2.435	0.0004314	0.00033	0.00049	0.0004	0.00035	0.0006347	9.648E-05	0.00037	0.000418	0.000305	0.000206	0.001238
2.442 (Glutaric acid/Unknown)	0.0005365	0.00038	0.00051	0.0004	0.000421	0.0007277	0.0001126	0.0004	0.000302	0.000387	0.000177	0.001318
2.469 (Glutamine)	0.0011368	0.01177	0.0006	0.0015	0.000445	0.0025377	0.0004433	0.00243	0.010217	0.003319	0.000967	0.001051
2.495 (3-OH Butyric acid)	0.0003039	0.00208	0.0001	0.001	0.000255	0.0007487	0.0003139	0.0015	0.000365	0.001408	0.000318	0.000172
2.512 (3-OH Butyric acid)	0.0005897	0.00136	0.00037	0.0012	0.000318	0.001036	0.0004018	0.00186	0.000448	0.00184	0.000508	0.000733
2.525	0.0002063	0.00024	0.00017	0.0001	0.000147	0.000344	6.382E-05	0.00023	0.00025	0.00039	0.000162	0.000658
2.532	0.0004641	0.00021	0.00044	0.0003	0.000279	0.0005204	1.908E-05	0.00032	0.000118	0.000265	0.000157	0.000811
2.547 (3-OH Butyric acid)	0.000791	0.00104	0.00071	0.0015	0.000574	0.0012228	0.0004035	0.00199	0.000315	0.001855	0.000566	0.001244

2.557 (3-OH Butyric acid)	0.0010212	0.00102	0.001	0.0017	0.000985	0.0016776	0.0009178	0.0023	0.000827	0.00306	0.000551	0.003124
2.588	0.013247	0.00553	0.01231	0.0106	0.011735	0.0155108	0.0049411	0.00989	0.004235	0.010576	0.008348	0.014599
2.647 (Methionine/Unknown)	0.0008775	0.0006	0.00078	0.0004	0.000306	0.0007687	0.0002125	0.00069	0.000757	0.000647	0.000564	0.000846
2.662 (Succinic acid)	0.0011413	0.00044	0.00157	0.0004	0.000487	0.0009784	0.0004032	0.0005	0.000439	0.000357	0.002015	0.001645
2.742 (Sarcosine)	0.0002421	7.1E-05	0.00012	0.0004	0.000925	0.0001256	0.0002572	0.00014	0.000101	0.000519	0.000216	0.000227
2.801	0.0004067	8.7E-06	-3.2E-06	-2E-06	8.85E-06	4.915E-05	-1.34E-05	6.3E-05	1.41E-05	0.000445	6.57E-05	-6.59E-06
2.831 (Citric acid)	0.0031747	0.00276	0.00275	0.0027	0.002731	0.0030483	0.0009833	0.00295	0.003179	0.003361	0.002124	0.003839
2.927 (Dimethylglycine)	0.0004855	0.00059	0.00053	0.0003	0.000225	0.0005533	0.0001262	0.0003	0.000598	0.000476	0.000262	0.000535
3.002 (Citric acid)	0.0108554	0.01095	0.01254	0.0094	0.01054	0.0124089	0.0028456	0.01044	0.011471	0.011912	0.006555	0.012708
3.052 (Creatine)	0.0046287	0.00439	0.00421	0.0025	0.00394	0.0065607	0.0016226	0.00334	0.004202	0.003204	0.002372	0.00378
3.105	0.0013087	0.00144	0.00055	0.0005	0.000446	0.0014438	0.0003652	0.00073	0.000614	0.001279	0.000814	0.00078
3.127 (Creatinine)	0.0030544	0.00444	0.00252	0.0024	0.002622	0.0035876	0.0013053	0.00187	0.003493	0.003038	0.002172	0.002815
3.141 (DMSO2)	0.0012457	0.00126	0.00081	0.0008	0.00077	0.0016196	0.000284	0.00089	0.00134	0.001604	0.001199	0.001238
3.188 (Acetylcarbitine/Choline)	0.0056501	0.00697	0.00777	0.0051	0.005132	0.0109731	0.0019391	0.00674	0.006809	0.006168	0.012622	0.006844
3.212 (Glucose)	0.0018251	0.00713	0.00213	0.0055	0.0042	0.0029772	0.0038149	0.00475	0.004805	0.008442	0.008509	0.001669
3.221 (Carnitine)	0.0015783	0.00209	0.00179	0.0016	0.002066	0.0026936	0.0012823	0.0021	0.001847	0.001745	0.002847	0.001605
3.239 (Glucose)	0.0078982	0.02458	0.00789	0.0192	0.014866	0.0135208	0.0122117	0.01623	0.021519	0.030533	0.028681	0.004201
3.262 (Betaine/Myo-inositol)	0.0075837	0.00788	0.00563	0.0039	0.005176	0.0133227	0.0026899	0.00385	0.008865	0.00619	0.004765	0.005176
3.283 (Myo-inositol)	0.0010216	0.00079	0.00049	0.0006	0.000562	0.0027343	0.0002536	0.00063	0.001192	0.000891	0.00062	0.00082
3.333	0.001089	0.00216	0.0012	0.0015	0.001981	0.0018532	0.0006457	0.0014	0.002192	0.001343	0.001484	0.001401
3.350 (Methanol)	0.0040955	0.00549	0.0048	0.0047	0.006092	0.0044532	0.001646	0.00422	0.005836	0.004971	0.003394	0.005637
3.394 Glucose)	0.0115531	0.04756	0.01344	0.0385	0.030958	0.0168668	0.0221811	0.03167	0.037862	0.061084	0.056698	0.008202
3.456 (Glucose)	0.0069544	0.04342	0.00923	0.0369	0.027614	0.0090052	0.0234279	0.02962	0.034016	0.060951	0.055698	0.003727
3.518 (Glucose)	0.010048	0.02565	0.00843	0.0202	0.01631	0.0231552	0.171171	0.0168	0.020786	0.032923	0.030317	0.005488
3.561	0.0019659	0.00383	0.00149	0.0024	0.00236	0.0026094	0.0031277	0.00218	0.002863	0.003092	0.003066	0.00132
3.599 (Myo-inositol/Unknown)	0.0057857	0.00522	0.00318	0.0031	0.002688	0.0160231	0.0011432	0.00284	0.004516	0.004214	0.004394	0.003139
3.647 (Ethanol)	0.0096536	0.01303	0.0097	0.012	0.012739	0.0149559	0.0047187	0.01146	0.011682	0.010522	0.016294	0.010663
3.728 (Glucose)	0.0213625	0.07122	0.02715	0.0562	0.047629	0.0275249	0.0370763	0.04794	0.056752	0.08823	0.078565	0.023024
3.783	0.0007175	0.00154	0.0005	0.001	0.000945	0.0007612	0.0023576	0.00102	0.001299	0.00165	0.002609	0.000568
3.801	0.0015116	0.00562	0.00145	0.0043	0.003523	0.0018605	0.0035636	0.00371	0.004457	0.007585	0.006971	0.001201

3.830 (Glucose)	0.0070094	0.02551	0.00895	0.0221	0.018643	0.0109771	0.0107972	0.0202	0.02309	0.036875	0.035068	0.008783
3.889 (Glucose)	0.0082416	0.03046	0.00836	0.0223	0.01996	0.0120168	0.3124096	0.01973	0.025324	0.036934	0.030497	0.012558
3.921	0.0018731	0.00214	0.00156	0.008	0.002506	0.0031167	0.0031402	0.00205	0.001669	0.001497	0.001787	0.001943
3.951 (Betaine)	0.0016468	0.00277	0.002	0.0012	0.001261	0.001665	0.0007959	0.00118	0.002439	0.001921	0.002094	0.002298
3.994	0.0058561	0.00614	0.00531	0.0039	0.004353	0.0072004	0.0017559	0.00457	0.005204	0.004256	0.005003	0.005488
4.051 (Myo-inositol)	0.0033972	0.00238	0.00225	0.0016	0.001581	0.0092086	0.0020131	0.00197	0.003005	0.002791	0.002596	0.003239
4.106 (Creatine)	0.0050575	0.005	0.00329	0.0028	0.003309	0.0060678	0.0031651	0.00383	0.005593	0.004019	0.002005	0.004249
4.144 (Proline/Unknown)	0.0007438	0.00107	0.00027	0.0038	0.000167	0.0004441	0.0001266	0.00056	0.000334	0.000419	0.002107	0.000373
4.195	0.0012672	0.00128	0.0014	0.0012	0.001086	0.0014505	0.0020548	0.00127	0.001327	0.001339	0.000624	0.001816
4.211	0.0008795	0.00106	0.00094	0.001	0.000675	0.0009409	0.0001703	0.00126	0.000551	0.001014	0.000773	0.000821
4.239	0.0010342	0.00194	0.00079	0.0028	0.000701	0.0017515	0.0008684	0.00398	0.001421	0.003979	0.00141	0.001139
4.287 (Creatinine)	0.0012307	0.00205	0.00138	0.0012	0.001331	0.0015606	0.0007097	0.00107	0.001917	0.001831	0.00091	0.001714
4.365 (Lactic acid)	0.1456819	0.0887	0.14156	0.1156	0.122838	0.1227152	0.0591713	0.12141	0.110302	0.075462	0.095326	0.13826
4.512	0.0003797	-0.0008	-0.00134	-0.0003	-0.00065	-0.000337	0.0001024	0.00014	0.000193	0.000138	0.000207	0.00018
4.632 (Glucose)	0.0004414	0.00852	-0.00133	0.0082	0.00497	-0.000216	0.0071045	0.00659	0.006803	0.014092	0.014316	-2.94E-05
5.169 (Mannose)	0.0002216	0.00048	-4.1E-05	0.0002	0.000213	1.491E-05	0.0002217	0.00042	0.000573	0.000648	0.000988	0.000328
5.221 (Glucose)	0.0013677	0.0126	0.00175	0.011	0.008022	0.0010167	0.0085085	0.00899	0.010429	0.019044	0.017522	0.000735
6.091 (Inosine)	8.27E-06	-2E-05	-0.00011	-0.0001	-0.0001	-0.000166	-1.79E-05	-1E-04	4.95E-05	-2.2E-05	-3.3E-05	-0.000111
6.890 (Tyrosine)	0.001739	0.00152	0.00157	0.0011	0.001501	0.0013361	0.0005338	0.00162	0.001849	0.001989	0.001155	0.002546
7.185 (Tyrosine)	0.0019634	0.00149	0.00182	0.0013	0.001632	0.0015204	0.0006246	0.00188	0.00208	0.002442	0.001222	0.002851
7.322 (Phenylalanine)	0.0018374	0.00168	0.00124	0.0014	0.00146	0.0013738	0.0007672	0.0018	0.002313	0.001503	0.001409	0.002314
7.407 (Phenylalanine/Histidine)	0.0041866	0.00385	0.0027	0.004	0.003651	0.0031394	0.0013743	0.00378	0.004983	0.003904	0.002821	0.004755
7.653	0.0005552	0.00034	0.00016	0.0001	0.000267	0.000284	3.994E-05	0.00028	0.00049	0.00032	0.000281	0.000473
7.792	0.0001352	-0.0002	-0.00014	-0.0001	-7.3E-05	-0.000104	8.237E-05	7.8E-05	0.00014	6.21E-05	6.6E-05	-8.12E-05
7.931	0.000241	2.9E-05	0.00016	1E-04	0.000111	0.0002199	9.57E-06	0.00016	0.00011	0.000119	7.63E-05	0.000309
8.224 (Hypoxanthine/Inosine)	0.0003619	0.00224	0.0004	0.001	0.000827	0.0005574	-1.44E-05	0.00037	0.001818	0.000717	-2.2E-05	0.001801
8.243 (Formic acid)	0.0017547	0.00053	0.00052	0.003	0.000543	0.0004488	0.0007459	0.00079	0.000571	0.001572	0.000318	0.001538
8.418 (Hypoxanthine)	0.0006108	0.00144	-0.00074	0.0003	0.000278	0.0001732	6.444E-05	0.00022	0.001838	0.000541	0.000298	0.00096
8.654 (Histidine)	0.0011809	0.00143	0.00082	0.0009	0.001379	0.0008033	0.0004564	0.00109	0.001511	0.001335	0.000711	0.001427

Table A1–3: ¹H-NMR spectral data for 114 bins relating to cord blood from 15 neonates not exposed to HIV.

Assignments: ppm and corresponding metabolite names	Control1	Control2	Control3	Control4	Control5	Control6	Control7	Control8	Control9	Control10	Control11	Control12	Control13	Control14	Control15
0.846	0.00082535	8.117E-05	0.0001452	1.28E-06	3.626E-05	-0.0004149	0.0050255	0.0001913	0.0001676	0.00013313	0.00024676	0.00012515	0.00015599	0.00004391	0.00024039
0.861	0.00018137	5.463E-05	6.001E-05	2.145E-05	5.736E-05	-3.019E-05	0.0010892	4.453E-05	5.559E-05	0.00005675	0.00007815	0.00004697	0.00006099	-4.9E-07	0.00006884
0.865	0.00015344	7.274E-05	9.232E-05	0.0001375	3.552E-05	1.206E-05	0.0006554	7.453E-05	7.225E-05	0.00005552	0.00014868	0.00005714	0.00012544	0.00002993	0.00013855
0.869 (2-OH Isovaleric acid)	0.00031814	0.0003213	0.000191	0.0001088	0.0001907	0.0001446	0.0016111	0.0002079	0.0001298	0.00023206	0.00012324	0.00012388	0.00012303	0.00007377	0.000169
0.889	0.00291957	0.0019281	0.0017111	0.0016177	0.0016005	0.0005428	0.013618	0.0029839	0.0014197	0.00094488	0.00214807	0.0012508	0.00170211	0.00108703	0.00178946
0.917 (Ethylmalonic acid)	0.00144314	0.0008907	0.0013155	0.0016263	0.0009368	0.0007815	0.0031275	0.0009896	0.0010078	0.00041132	0.00164468	0.00114079	0.0012599	0.00070941	0.00139452
0.925	0.00059596	0.0012666	0.0006098	0.0007642	0.000397	0.0003229	0.0019334	0.0006471	0.0007265	0.0002639	0.00077367	0.00038085	0.0005186	0.00064315	0.00056907
0.939 (Isoleucine/Leucine)	0.00631515	0.0106506	0.0071636	0.0064909	0.0067035	0.0080319	0.0096375	0.0078887	0.0067607	0.00362195	0.0072205	0.00656437	0.00516095	0.00613286	0.00675533
0.960 (Leucine)	0.00772696	0.0093469	0.0074821	0.0060062	0.0070403	0.0087276	0.0088581	0.0075978	0.0075595	0.00313816	0.00718716	0.00791942	0.00456791	0.00556172	0.00671531
1.004 (Isoleucine/Valine)	0.01163836	0.0125552	0.0128544	0.0097433	0.0128711	0.0143918	0.0111132	0.0123813	0.0113615	0.00491254	0.01122215	0.01432353	0.00782851	0.00812264	0.01185989
1.014 (Valine)	0.00795517	0.0080753	0.0087679	0.006715	0.0090255	0.0099278	0.0068293	0.0088955	0.0079577	0.00332868	0.00755986	0.01036659	0.00534526	0.00523505	0.00812638
1.065	0.00022966	0.0002684	0.0002593	0.0002072	0.0003523	0.000297	0.0004486	0.0002566	0.0002599	0.00010844	0.00034404	0.00017749	0.00028694	0.00041148	0.00028742
1.080	0.00052037	0.0005913	0.0007814	0.0008751	0.0008088	0.0006793	0.0008963	0.0005172	0.0006298	0.00023089	0.00095274	0.00047876	0.00103289	0.00083914	0.00085269
1.094	0.00033165	0.0002838	0.0006237	0.0006554	0.0004677	0.0001979	0.000719	0.0003618	0.0003975	0.00015021	0.00066856	0.00028597	0.00064755	0.00044928	0.00062124
1.104	0.00034642	-2.687E-05	0.0001809	0.0005155	-4.64E-06	-0.0001252	0.0003468	1.605E-05	0.0001228	0.00003327	0.00051792	0.00019531	0.00034232	0.00005597	0.00020548
1.113	0.0000992	0.0003605	0.0002037	0.0001064	0.0002013	7.878E-05	0.0003758	0.0002624	0.0001278	0.00018903	0.00011978	0.00005326	0.00010819	0.00014868	0.00012216
1.122	0.00083288	0.0008987	0.0010317	0.0012027	0.0008442	0.000474	0.0012393	0.0008419	0.0008847	0.00173844	0.00129389	0.00080918	0.00119998	0.00070506	0.00088828
1.130 (Propylene glycol)	0.00006357	0.000804	0.0008185	0.0003348	0.0006556	0.0005459	0.0004939	0.0007836	0.0003837	0.000402	0.00019917	0.00013755	0.00038259	0.00038917	0.00044094
1.136	0.00037863	0.0007478	0.0007594	0.0006658	0.0007609	0.0003841	0.0010282	0.0006986	0.0007097	0.00157976	0.00067599	0.00051985	0.0007392	0.00058931	0.00053331
1.144 (Propylene glycol)	0.00015079	0.0006728	0.0008458	0.0003585	0.0005942	0.0003973	0.0007788	0.0007369	0.0003959	0.00033271	0.00020548	0.00008844	0.00043189	0.00038064	0.00050081
1.170 (Ethanol)	0.00187931	0.0123677	0.0101486	0.0087864	0.0145792	0.0138578	0.0113998	0.0169995	0.0095501	0.00401197	0.00965984	0.01982553	0.00892859	0.00897458	0.00820451
1.230 (3-OH-Butyric acid)	0.00726902	0.0199972	0.0243801	0.0235944	0.0155316	0.0088979	0.0096523	0.029852	0.0105929	0.00457317	0.01196317	0.00475191	0.00417295	0.00419566	0.00240145
1.275 (C13 Lactic acid)	0.00211428	0.0022751	0.002873	0.0025739	0.0025071	0.0007715	0.0050723	0.0019039	0.0021962	0.00159921	0.00313744	0.0021163	0.00355167	0.00230845	0.00336251
1.310	0.00022945	0.0002846	0.0004206	0.0002094	0.0003531	-0.000109	0.0010471	0.000442	0.0002985	0.00050557	0.00043671	0.00032812	0.00033969	0.00030155	0.00036489
1.321 (Threonine/3-hydroxyisovaleric acid)	0.00384299	0.0041468	0.005935	0.0042437	0.0074497	0.0039564	0.0060501	0.0044126	0.003382	0.00379326	0.00415584	0.00559998	0.00349394	0.00304371	0.00397791
1.334 (Threonine)	0.00379668	0.0035042	0.0056892	0.0040601	0.0073684	0.0037805	0.0059791	0.0042269	0.0033492	0.0035655	0.00399405	0.0054862	0.00325948	0.00278498	0.00391495

1.357	0.00032787	0.0001905	0.0005778	0.0002058	0.0003476	-0.0003163	0.001396	0.0003411	0.0002793	0.00071048	0.00044289	0.00017431	0.00031096	0.00018202	0.00031851
1.408 (Lactic acid)	0.27501192	0.3276535	0.4061555	0.41187	0.3749128	0.1613247	0.3005436	0.2172322	0.2818534	0.19437982	0.45089217	0.30389616	0.55133178	0.31068023	0.49604532
1.468	0.00113395	0.0015284	0.0008365	0.0009209	0.0008812	0.0010809	0.0026227	0.0010765	0.0009467	0.00054505	0.00094351	0.00101192	0.00093713	0.00071631	0.00121748
1.507 (Alanine)	0.01957353	0.0262919	0.0241105	0.0262651	0.0265485	0.0212315	0.0215128	0.0202773	0.0176192	0.01231175	0.02916118	0.02264894	0.02315373	0.01586811	0.02439065
1.536 (C13 Lactic)	0.00260962	0.0028199	0.0027536	0.0028212	0.0026895	0.0015621	0.0046693	0.0021811	0.0021289	0.00117908	0.00287892	0.00220298	0.00312211	0.00187983	0.00332077
1.561 (Pyruvic acid)	0.00242867	0.0003374	0.0002224	0.000785	9.623E-05	3.412E-05	0.0021007	0.0001714	0.0002332	0.00020999	0.00291639	0.00048618	0.00129209	0.00012687	0.00044756
1.722 (Lysine)	0.0174771	0.0173652	0.0163307	0.0145976	0.0166581	0.0143913	0.0294167	0.0171917	0.0161566	0.0071008	0.01495357	0.01634353	0.01067236	0.010578	0.01540708
1.916 (Lysine/Glutaric acid)	0.01589955	0.014307	0.0167214	0.0144476	0.0166218	0.0120431	0.0246182	0.0156459	0.0144909	0.00724645	0.0148784	0.015553021	0.01174054	0.00937049	0.01527905
2.024 (Proline)	0.00787284	0.0073778	0.0071692	0.0057292	0.0065232	0.00575	0.0175841	0.0061709	0.006422	0.00487171	0.00740283	0.00740687	0.0057248	0.00548761	0.00662766
2.080 (Acetic acid)	0.00297968	0.0108105	0.0099295	0.0117144	0.0109292	0.0134837	0.0136308	0.0099843	0.0112807	0.00658614	0.0093166	0.00322785	0.01366106	0.01355742	0.01395971
2.124 (Methionine/Acetylcarnitine)	0.00140896	0.0005627	0.0014302	0.0011844	0.0012246	0.0008627	0.001419	0.0008834	0.0009465	0.00026924	0.00129846	0.00164751	0.00088131	0.00058474	0.00112002
2.171 (Glutamine)	0.01632972	0.0134456	0.0160683	0.0132527	0.0144951	0.0153495	0.0185447	0.0126481	0.0112885	0.00897787	0.01932389	0.01994651	0.01401801	0.01048386	0.01474585
2.219 (Acetone)	0.00232725	0.000733	0.0016647	0.0017599	0.0006672	0.0007022	0.0012159	0.0010623	0.0008421	0.00043557	0.0004982	0.00177069	0.00056447	0.00037715	0.00063721
2.332 (Proline/Unknown)	0.00380942	0.0037913	0.0038406	0.002992	0.0034415	0.0029212	0.0063912	0.003526	0.0033624	0.00227832	0.00374607	0.00391119	0.00225135	0.00277066	0.00335381
2.378 (Pyruvic acid)	0.00385012	0.0007119	0.0010088	0.001888	0.0008185	0.000652	0.0011847	0.0005947	0.0006793	0.00063276	0.00624521	0.00129897	0.00300506	0.00075982	0.00137225
2.396	0.00022658	0.0002539	0.0002672	0.0002135	0.0003008	0.000372	0.0003751	0.0002733	0.0002837	0.00018234	0.0004423	0.00028502	0.00033896	0.00047268	0.00030927
2.402	0.00037079	0.0005544	0.0005857	0.0003807	0.0005719	0.0002147	0.0006665	0.000211	0.0002989	0.00019517	0.00092205	0.00025747	0.0004462	0.00024076	0.00059952
2.410	0.00021076	0.0002536	0.0003455	0.0002734	0.0002981	0.0002214	0.0004309	0.0002594	0.0003064	0.00021473	0.00097865	0.00030576	0.0003753	0.00030091	0.0005568
2.415 (Glutaric acid/Unknown)	0.00025523	0.000477	0.0005333	0.0002748	0.0004395	0.0002324	0.0004459	0.0001814	0.0002207	0.00015114	0.00072321	0.00024643	0.00033278	0.00020144	0.00042196
2.423	0.00063904	0.000845	0.0010354	0.0006706	0.0009203	0.0006372	0.0007801	0.0004658	0.0006292	0.00041422	0.00207507	0.00069284	0.00096435	0.00052876	0.00125585
2.429 (Glutaric acid)	0.00023579	0.0004591	0.0004765	0.000287	0.0004883	0.0002476	0.0004582	0.0002237	0.0002549	0.00016308	0.0007187	0.00031883	0.00034109	0.00021751	0.00045967
2.435	0.00037069	0.0004028	0.0004594	0.000278	0.0004358	0.0002434	0.000481	0.0002107	0.0002855	0.00017595	0.00087846	0.00047469	0.0003821	0.00022436	0.00055172
2.442 (Glutaric acid/Unknown)	0.00029689	0.0004038	0.0005163	0.0003401	0.0004037	0.0002807	0.00055	0.0002661	0.000358	0.00021547	0.00106538	0.00037386	0.00044096	0.00027339	0.00059719
2.469 (Glutamine)	0.01040785	0.003569	0.0053658	0.0030168	0.0021277	0.0007174	0.0045789	0.0037651	0.0027672	0.00093068	0.00362547	0.01230866	0.00087324	0.00089473	0.000853
2.495 (3-OH Butyric acid)	0.00183163	0.0022998	0.0028346	0.0025257	0.0015853	0.0009461	0.0009536	0.0031899	0.0011932	0.00050197	0.00132641	0.0005052	0.00037938	0.00040695	0.00026372
2.512 (3-OH Butyric acid)	0.00127369	0.0025979	0.0031964	0.0028834	0.0018709	0.0011257	0.0012938	0.0035767	0.0015078	0.000653	0.00183804	0.00065078	0.0005499	0.00055465	0.00046171
2.525	0.00027119	9.936E-05	0.0001767	0.0001666	0.0001101	7.02E-06	0.0004555	0.0002386	0.0003152	0.00020285	0.00060982	0.00029684	0.0002123	0.00019793	0.00032719
2.532	0.00018795	0.0003443	0.0004652	0.0002962	0.0004276	0.0001534	0.0005269	0.000176	0.0002335	0.00014693	0.00063374	0.00012916	0.00028266	0.00017112	0.00038445
2.547 (3-OH Butyric acid)	0.00102739	0.0028848	0.0033753	0.003154	0.0022878	0.0013311	0.0013978	0.0036284	0.0015046	0.00076621	0.00232867	0.00046819	0.00075015	0.00065982	0.0007794
2.557 (3-OH Butyric acid)	0.00114387	0.0030992	0.0035657	0.0034371	0.002533	0.001801	0.0013259	0.0047613	0.0022525	0.00176243	0.00386934	0.00106022	0.00168541	0.00171587	0.00177955

2.588	0.00418423	0.0125936	0.013272	0.0131364	0.013619	0.0131782	0.0137563	0.0132519	0.0083821	0.00760085	0.01336594	0.00511953	0.00971739	0.00855346	0.01134871
2.647 (Methionine/Unknown)	0.00078714	0.0003858	0.0006079	0.0005302	0.0004953	-1.16E-06	0.0008728	0.0005137	0.0006261	0.00024087	0.00104659	0.00077019	0.00040887	0.00033924	0.00073559
2.662 (Succinic acid)	0.0002932	0.0006688	0.0005803	0.000631	0.0002806	0.0001937	0.0008645	0.0003848	0.0004283	0.0006535	0.00137372	0.00025764	0.00082635	0.00062961	0.00081259
2.742 (Sarcosine)	0.00007531	0.0007421	0.0001979	0.0002839	0.0004183	0.0005043	0.0005115	0.0003477	0.0003263	0.00049916	0.00015806	0.00005245	0.00013884	0.00031424	0.00024572
2.801	0.00003516	-4.909E-05	-2.04E-06	-1.31E-05	-2.252E-05	-0.0001555	0.0001903	8.74E-06	5.481E-05	0.00006808	0.00017048	0.00004037	-3.748E-05	-3.134E-05	-1.455E-05
2.831 (Citric acid)	0.00327694	0.0026507	0.0026337	0.0038772	0.0031318	0.0010461	0.0046221	0.0035541	0.0021581	0.00296953	0.00383453	0.0039576	0.00277118	0.00305158	0.0029591
2.927 (Dimethylglycine)	0.00060376	0.000284	0.0004841	0.0003906	0.0007675	0.0059374	0.0012065	0.0002263	0.0003666	0.00065464	0.0004271	0.00055288	0.00037783	0.00027874	0.00049496
3.002 (Citric acid)	0.01213753	0.0115537	0.0130124	0.0118413	0.0135749	0.0088517	0.0179597	0.0126056	0.0106039	0.00673183	0.01086479	0.01230449	0.00937741	0.00845086	0.01127788
3.052 (Creatine)	0.00283651	0.0038886	0.003656	0.0034857	0.0032942	0.0021775	0.0032916	0.0031635	0.0028772	0.00214894	0.00362119	0.00337645	0.00259699	0.00269269	0.00376036
3.105	0.00062412	0.0006023	0.0006162	0.0011797	0.0007832	-5.988E-05	0.0016862	0.0005674	0.0007202	0.00077156	0.00093744	0.00061419	0.00044798	0.00042752	0.00058909
3.127 (Creatinine)	0.00222478	0.0028321	0.0024488	0.0022952	0.0035165	0.0033716	0.0024971	0.0023849	0.0022951	0.00134629	0.00402383	0.00340964	0.0024939	0.0024932	0.00279608
3.141 (DMSO2)	0.00099014	0.0008288	0.0007753	0.0008838	0.0010842	0.0005109	0.0015272	0.0009728	0.0008778	0.00075229	0.00147131	0.00143129	0.00078324	0.00079714	0.00103586
3.188 (Acetylcarnitine/Choline)	0.00418101	0.0075193	0.0052523	0.0054735	0.0078787	0.0040895	0.0066653	0.0046715	0.0049854	0.00535387	0.00599255	0.00751797	0.00385849	0.00359578	0.0079493
3.212 (Glucose)	0.00733188	0.0058627	0.0037021	0.0033659	0.0039805	0.0107468	0.0049266	0.0081733	0.0079104	0.00436056	0.00172965	0.00544754	0.00100361	0.00819744	0.00138669
3.221 (Carnitine)	0.00176758	0.0024938	0.0018509	0.0014206	0.0020877	0.0026966	0.0016629	0.0023676	0.0024607	0.00094617	0.00359685	0.00209009	0.00149307	0.00196583	0.00213575
3.239 (Glucose)	0.02650284	0.0207371	0.0115565	0.0124672	0.0147447	0.0388407	0.0155719	0.0289003	0.0281635	0.01503396	0.00745586	0.019538	0.00356019	0.02983874	0.00581451
3.262 (Betaine/Myo-inositol)	0.00579802	0.0062717	0.0055906	0.0051871	0.0055877	0.0042858	0.0035766	0.0051003	0.0045162	0.0029576	0.00651028	0.00734837	0.0050831	0.00391133	0.00640963
3.283 (Myo-inositol)	0.000663	0.0005286	0.0004849	0.0004926	0.0005083	0.0006283	0.0007298	0.0006892	0.0007733	0.00068102	0.001348	0.00107245	0.00064281	0.00078042	0.00077126
3.333	0.00165346	0.0010315	0.0015881	0.0014017	0.0014763	0.0018751	0.0013035	0.0014079	0.001469	0.00092973	0.00156977	0.00245161	0.00133627	0.00132421	0.00140758
3.350 (Methanol)	0.00834345	0.0045347	0.0060132	0.0043688	0.0065295	0.0060009	0.0042315	0.0048472	0.0047406	0.00287315	0.00498327	0.00605095	0.00408131	0.00430663	0.00478178
3.394 Glucose)	0.05979057	0.0385853	0.0216719	0.0243322	0.0284588	0.0773946	0.0291939	0.0576489	0.0564875	0.03003086	0.00975572	0.03849633	0.00579774	0.05962807	0.00898875
3.456 (Glucose)	0.0522593	0.0364507	0.0177894	0.0227759	0.0251551	0.0810748	0.028067	0.0590848	0.0580402	0.03030735	0.00480448	0.03753712	0.00263526	0.06057827	0.00546236
3.518 (Glucose)	0.02940011	0.0217663	0.0113282	0.0138415	0.0165057	0.0406396	0.0172589	0.0302054	0.0296366	0.02010317	0.00913799	0.02262585	0.0038733	0.03108918	0.00650144
3.561	0.00675518	0.0027147	0.0018905	0.0021941	0.0031392	0.002679	0.002602	0.0020168	0.0018648	0.00328785	0.00177669	0.00305358	0.00134827	0.00179419	0.00172656
3.599 (Myo-inositol/Unknown)	0.00766255	0.0037068	0.0024425	0.0024519	0.0029639	0.002737	0.0035324	0.0035901	0.0027901	0.00546755	0.00624907	0.00403062	0.00247204	0.0037372	0.00271073
3.647 (Ethanol)	0.01129954	0.0140289	0.0107571	0.0101996	0.0155954	0.0135933	0.0105071	0.0157497	0.0106423	0.01077975	0.01087413	0.01819774	0.0090666	0.01048508	0.00919292
3.728 (Glucose)	0.07843954	0.0581439	0.0376762	0.0396139	0.0484696	0.1109606	0.0478718	0.084955	0.0806922	0.40399155	0.01802116	0.0645068	0.01530729	0.08057445	0.01891318
3.783	0.00125009	0.0012039	0.0007161	0.0008068	0.0010171	0.0019423	0.0012788	0.0013018	0.0013095	0.00124859	0.00061268	0.00142681	0.00049929	0.0013412	0.00075933
3.801	0.00594135	0.0044375	0.0023158	0.0029963	0.0032896	0.0086779	0.0039878	0.006553	0.0064824	0.00503243	0.00135433	0.00529203	0.0008005	0.0071437	0.0011456
3.830 (Glucose)	0.03047943	0.0215251	0.0114076	0.017081	0.0180128	0.0477443	0.0187438	0.0359556	0.0347166	0.01819883	0.00837481	0.02706206	0.00654538	0.03549321	0.00838047

3.889 (Glucose)	0.03344549	0.0223695	0.0142594	0.015967	0.0161462	0.048928	0.0192902	0.0358918	0.0338917	0.02259811	0.01204883	0.02964537	0.00711461	0.03443387	0.00956054
3.921	0.00105244	0.002692	0.0027154	0.002209	0.0018454	0.0034534	0.003545	0.0018737	0.0014997	0.00104621	0.001818	0.00197548	0.00230178	0.00143487	0.00197725
3.951 (Betaine)	0.00429463	0.001739	0.0022449	0.0013054	0.0026801	0.0018375	0.0020299	0.0018562	0.0019014	0.0006868	0.00198544	0.00285335	0.00190143	0.00128354	0.00225117
3.994	0.00456606	0.0053457	0.007279	0.0034497	0.0062067	0.0059457	0.0049262	0.0042937	0.0040204	0.00250329	0.00458684	0.00630108	0.00451619	0.00379669	0.00484325
4.051 (Myo-inositol)	0.00185203	0.0021922	0.0012398	0.0016914	0.0018183	0.0016607	0.0026186	0.0018297	0.0018752	0.00213858	0.00417082	0.00272999	0.00185685	0.00228848	0.00250621
4.106 (Creatine)	0.00330334	0.0037062	0.0031383	0.0031453	0.0025339	0.0029539	0.0054194	0.0035949	0.0030758	0.00284791	0.00433447	0.00562267	0.00252849	0.00294901	0.00364683
4.144 (Proline/Unknown)	0.0007113	0.000929	0.0009951	0.0003224	0.0003118	9.558E-05	0.0022974	0.0002473	0.000311	0.00033375	0.00057918	0.00024929	0.00030312	0.000102	0.00038105
4.195	0.00123661	0.0010825	0.001416	0.0014009	0.0012719	0.0012762	0.0019977	0.0012708	0.0011934	0.00066629	0.00133062	0.00155199	0.00108476	0.00085638	0.00118933
4.211	0.00098329	0.0014996	0.0019008	0.001565	0.0012485	0.0006961	0.0016909	0.0015213	0.0009129	0.00039921	0.00115246	0.00065355	0.00068875	0.00049147	0.00077626
4.239	0.00229077	0.0054999	0.005937	0.0060818	0.0035262	0.0022636	0.0043443	0.0074089	0.0035008	0.00150335	0.00421491	0.00130692	0.0015594	0.00184334	0.00168421
4.287 (Creatinine)	0.00112544	0.0011691	0.0012062	0.0011841	0.0016392	0.0016859	0.0021177	0.0012497	0.0013876	0.0009944	0.00255042	0.00184692	0.00136503	0.00124817	0.00155193
4.365 (Lactic acid)	0.08165197	0.0981044	0.1223421	0.1229276	0.1121036	0.049409	0.0906414	0.0674417	0.0863134	0.0609093	0.14004347	0.09273973	0.15913968	0.08894014	0.15298693
4.512	-0.0003164	-0.00101	-0.0013132	-0.000394	-0.0012907	-0.0001422	0.0001078	1.93E-06	0.0002762	0.0002608	0.00035258	0.00025337	-0.0001171	0.00000103	0.0003352
4.632 (Glucose)	0.01264338	0.0065525	0.0012944	0.0045173	0.003243	0.0223454	0.006231	0.0138322	0.0146173	0.00611387	0.00016771	0.00892078	-0.0003524	0.01570051	0.00090274
5.169 (Mannose)	0.00033398	0.0006293	0.0000858	0.0002098	0.0001079	0.0004635	0.0002542	0.0008184	0.0003371	0.0003514	-1.317E-05	0.00051663	0.00088823	0.00108673	0.00037883
5.221 (Glucose)	0.0154459	0.0109806	0.0047751	0.006466	0.0066942	0.0247629	0.0078733	0.0191073	0.0176952	0.00924114	0.00004945	0.0116056	0.00194197	0.01856591	0.00145155
6.091 (Inosine)	-8.078E-05	-0.0001092	-0.0001363	-6.22E-05	-0.0001095	-0.0005302	-0.0001412	-0.00014	-0.000276	-0.000108	-0.0002887	0.00076154	-0.0003148	-0.0002007	-0.0001088
6.890 (Tyrosine)	0.0018524	0.0014489	0.0015589	0.0014632	0.001217	0.0009334	0.00221	0.0016783	0.0014009	0.00093352	0.00131267	0.00176988	0.00155296	0.00101437	0.00142837
7.185 (Tyrosine)	0.00214001	0.001674	0.0018537	0.0016361	0.0016678	0.0011021	0.0036268	0.0019455	0.0016385	0.00116423	0.00170159	0.00224982	0.00154089	0.00122097	0.00179809
7.322 (Phenylalanine)	0.00206537	0.0014718	0.0012221	0.0013711	0.0013796	0.0008171	0.0023805	0.0018455	0.0010785	0.0009807	0.00184999	0.00232461	0.00133782	0.00119962	0.00164022
7.407 (Phenylalanine/Histidine)	0.00399813	0.0031119	0.0033208	0.003041	0.0034554	0.0025344	0.0042669	0.0040679	0.0025836	0.00202209	0.00363187	0.00481485	0.00250429	0.0026235	0.00358649
7.653	0.00040446	0.0001732	0.0002586	0.0001472	2.706E-05	-0.0003714	0.0009033	0.0003255	9.923E-05	0.00018683	0.00024043	0.0003836	0.00024566	0.00007425	0.00033034
7.792	0.00004305	-0.0001536	-8.649E-05	-0.000182	-0.0003194	-0.0005415	0.0005954	3.328E-05	-0.00013	0.00015038	-0.0001624	0.00000928	-0.0001737	0.00000828	-3.61E-06
7.931	0.00013124	0.000138	0.0001533	7.975E-05	4.495E-05	-9.871E-05	0.000492	5.667E-05	6.505E-05	0.00013813	0.00019679	0.00006575	0.00015942	0.00008793	0.0002152
8.224 (Hypoxanthine/Inosine)	0.0005016	0.000282	0.0005369	0.0001862	0.0009455	-0.0001293	0.001183	0.0003617	0.000207	0.00025479	0.00073307	0.00408254	0.00031041	0.00043976	0.00010495
8.243 (Formic acid)	0.00036356	0.0006056	0.0006535	0.0009795	0.0007954	0.0012275	0.0016123	0.0009679	0.0007998	0.00078414	0.00039112	0.00056289	0.00072858	0.00108817	0.0003568
8.418 (Hypoxanthine)	0.00127625	-0.0003382	0.0002007	-0.000707	8.977E-05	-0.0020171	0.0050952	-0.000258	0.0001077	0.00004007	0.00092132	0.00415207	-0.001187	-0.0007268	0.00047455
8.654 (Histidine)	0.00114976	0.0008005	0.0012701	0.0007529	0.0013576	0.0011508	0.001098	0.0010493	0.0007787	0.0005136	0.00104681	0.00138676	0.00047951	0.00066598	0.00092562

Table A1–4: ¹H-NMR spectral data for 114 bins relating to plasma from 10 infants.

Assignments: ppm and corresponding metabolite names	Infant1	Infant2	Infant3	Infant4	Infant5	Infant6	Infant7	Infant8	Infant9	Infant10
0.846	0.00027951	0.00016191	0.00024889	0.00046849	0.00067898	0.03560022	0.00043826	0.00012864	0.00012924	-0.00001623
0.861	0.00008423	0.00005097	0.0000767	0.00036295	0.00006375	0.00023634	0.00011342	0.00032333	0.00059905	0.00003333
0.865	0.00013782	0.00009851	0.00017305	0.00011686	0.00007843	0.00015791	0.00020846	0.00014688	0.00010993	0.00010168
0.869 (2-OH Isovaleric acid)	0.00017429	0.00008394	0.00011715	0.00027082	0.00010053	0.00026805	0.00026621	0.0001549	0.00007274	0.00007772
0.889	0.00537668	0.0038803	0.00186194	0.01021679	0.00218102	0.00425097	0.00708809	0.00558321	0.00468613	0.00403452
0.917 (Ethylmalonic acid)	0.00119525	0.00152653	0.00072821	0.00124652	0.00095517	0.00107556	0.00138189	0.00113497	0.00093096	0.00129612
0.925	0.0003705	0.00048123	0.00055824	0.00059599	0.00050076	0.00079937	0.00044374	0.00043322	0.00013777	0.00040063
0.939 (Isoleucine/Leucine)	0.0057914	0.00967061	0.00569842	0.00817039	0.00440117	0.00695789	0.00672971	0.00559287	0.00510502	0.00784364
0.960 (Leucine)	0.00825343	0.01357593	0.00858126	0.01111284	0.00468385	0.00826819	0.00851609	0.00701059	0.00805801	0.01215104
1.004 (Isoleucine/Valine)	0.00804544	0.01893126	0.01128987	0.01014429	0.00719151	0.01063654	0.01188693	0.01090303	0.00897459	0.0180289
1.014 (Valine)	0.0053785	0.01324932	0.00751551	0.00658334	0.00498971	0.00701116	0.00835926	0.00755221	0.00624023	0.0131972
1.065	0.00008248	0.00007548	0.00015805	0.00017277	0.00007298	0.00019029	0.00028515	0.00014271	0.0000581	0.00008769
1.080	0.00035195	0.00029663	0.00043108	0.00045686	0.00027357	0.00053632	0.00047386	0.00039828	0.00023288	0.00044907
1.094	0.00013069	0.00017475	0.00027437	0.00023458	0.00020636	0.00040681	0.00030857	0.00020851	0.00015664	0.00025202
1.104	0.00010371	0.00016922	0.00020797	0.00002658	0.00022364	0.00024906	0.00007852	0.00026498	0.00015072	0.00024649
1.113	0.00006705	0.00003485	0.00005855	0.00010364	0.00004405	0.00024963	0.00010975	0.00002372	-0.00000791	-0.0000226
1.122	0.00305215	0.00065651	0.00061322	0.00040426	0.00059826	0.0096565	0.00095466	0.00129975	0.00072847	0.00090941
1.130 (Propylene glycol)	0.00060763	0.00013519	0.00012953	0.0004872	0.00015766	0.00094171	0.00052114	0.00015394	0.00010803	0.00025993
1.136	0.00249496	0.00035874	0.00039812	0.00040594	0.0003736	0.00889213	0.00084003	0.00104979	0.00041084	0.00049293
1.144 (Propylene glycol)	0.0004267	0.00019046	0.00011837	0.00065164	0.00015224	0.00079244	0.00059745	0.00014445	0.00004855	0.0002315
1.170 (Ethanol)	0.00148178	0.0009497	0.00077027	0.00143953	0.00216228	0.0197421	0.00257406	0.00423814	0.00039082	0.00026342
1.230 (3-OH-Butyric acid)	0.00161356	0.00168678	0.0015928	0.01096324	0.00080986	0.03700093	0.00228934	0.00141052	0.00946701	0.00208671
1.275 (C13 Lactic acid)	0.00183258	0.0012143	0.00103898	0.00182122	0.00101777	0.00182075	0.00177059	0.0020917	0.00129902	0.00062891
1.310	0.00078309	0.00044797	0.00026794	0.00099566	0.0003108	0.0034695	0.00086519	0.00071762	0.00040282	0.0005154
1.321 (Threonine/3-hydroxyisovaleric acid)	0.00655725	0.00473388	0.00274688	0.00391029	0.0017318	0.01266707	0.00375248	0.00398433	0.00394523	0.00501643
1.334 (Threonine)	0.00523978	0.00449086	0.00238309	0.00331685	0.00139753	0.00481116	0.00313561	0.00353715	0.00379197	0.00461254
1.357	0.00049553	0.00025971	0.00008149	0.00081481	0.00019357	0.00050402	0.00073226	0.00040021	0.00017776	0.0003463
1.408 (Lactic acid)	0.13939042	0.08497244	0.11651724	0.10780101	0.09255465	0.10201331	0.07278352	0.19343242	0.07813467	0.07310981
1.468	0.00061714	0.0008765	0.00097791	0.00096216	0.00040257	0.00069048	0.00078679	0.0005149	0.00026236	0.00030553
1.507 (Alanine)	0.02358948	0.02501341	0.02281264	0.02360456	0.00976217	0.0134787	0.02908181	0.02439938	0.02836946	0.03083301
1.536 (C13 Lactic)	0.00201065	0.00174276	0.00234483	0.00268601	0.0010264	0.00133663	0.00209177	0.00205366	0.00110999	0.0012463
1.561 (Pyruvic acid)	0.00071325	0.0010285	0.00128558	0.0004301	0.0004428	0.00177048	0.00069858	0.00144177	0.00066662	0.0009791
1.722 (Lysine)	0.0152969	0.01979419	0.01615445	0.01725525	0.00785421	0.01310063	0.01315649	0.0103775	0.00890241	0.0139498
1.916 (Lysine/Glutaric acid)	0.01409766	0.01607359	0.01442521	0.01490351	0.0073636	0.01191833	0.01274615	0.00860445	0.0062663	0.01019028
2.024 (Proline)	0.00782644	0.01153407	0.00977418	0.00946106	0.004658	0.00605406	0.01269384	0.00945406	0.00551295	0.0082402
2.080 (Acetic acid)	0.01227633	0.01200747	0.01015852	0.01115592	0.00590056	0.01718154	0.02913243	0.01136724	0.01492276	0.0196677
2.124 (Methionine/Acetylcarnitine)	0.00121529	0.0017727	0.00176298	0.00143136	0.0008415	0.00104771	0.00135018	0.00116353	0.00127934	0.0016176

2.171 (Glutamine)	0.02292556	0.0210365	0.02230327	0.02033586	0.00931962	0.01337116	0.02000688	0.01357321	0.0149562	0.01786755
2.219 (Acetone)	0.00120345	0.00132566	0.00127741	0.00324752	0.00076178	0.00797458	0.0010703	0.00100483	0.00422925	0.00134599
2.332 (Proline/Unknown)	0.00316936	0.00571578	0.00467344	0.00456784	0.00247668	0.00790231	0.00608315	0.00450402	0.00268217	0.00336035
2.378 (Pyruvic acid)	0.00166379	0.00228741	0.0011285	0.00076997	0.00092487	0.00273209	0.00147184	0.0032266	0.00213306	0.00335742
2.396	0.00024871	0.0001817	0.00024781	0.00020472	0.00014583	0.00044215	0.00020051	0.00026175	0.00015794	0.00027476
2.402	0.00033492	0.00023631	0.00033268	0.00057453	0.00030872	0.0010082	0.00070747	0.00011155	0.00025332	0.00029423
2.410	0.0004432	0.00018748	0.00017668	0.00040882	0.00023942	0.00038066	0.00029382	0.00020524	0.00031298	0.0003883
2.415 (Glutaric acid/Unknown)	0.00038249	0.00019659	0.00029094	0.00038846	0.00019848	0.00049322	0.00049924	0.00015179	0.00014442	0.00021674
2.423	0.00080205	0.00063098	0.00057758	0.00116922	0.00058755	0.00081282	0.00113708	0.00044951	0.00059732	0.0008174
2.429 (Glutaric acid)	0.00040754	0.00020642	0.00046777	0.00038821	0.00020378	0.00035571	0.00052561	0.00019042	0.00010659	0.0002066
2.435	0.00059528	0.00049208	0.00029464	0.0005668	0.00028795	0.00042703	0.00058688	0.00036268	0.00036267	0.0005143
2.442 (Glutaric acid/Unknown)	0.00051364	0.00028615	0.00053043	0.00055609	0.00027498	0.00057447	0.00050917	0.00023967	0.000098	0.00026323
2.469 (Glutamine)	0.0083428	0.01280877	0.01172567	0.00966091	0.00208599	0.00822468	0.00856664	0.00787646	0.00917213	0.01271446
2.495 (3-OH Butyric acid)	0.00050596	0.00087052	0.00043835	0.0020679	0.00025189	0.00425971	0.0011841	0.00025192	0.00142165	0.00018063
2.512 (3-OH Butyric acid)	0.00044329	0.00048921	0.00065817	0.0016993	0.00023444	0.00448124	0.00074502	0.00029312	0.00126093	0.00028852
2.525	0.00037444	0.00024139	0.00012493	0.00033823	0.0001309	0.00015806	0.00016149	0.00016769	0.00018759	0.00024301
2.532	0.00022599	0.00009355	0.00033296	0.00040702	0.00022604	0.00021987	0.00040754	0.00008322	0.00006166	0.00002494
2.547 (3-OH Butyric acid)	0.0004888	0.00031085	0.00049978	0.00165881	0.00036296	0.00451491	0.00063652	0.0002403	0.00118356	0.00015543
2.557 (3-OH Butyric acid)	0.00167144	0.00051513	0.00048587	0.00184099	0.0004966	0.00457063	0.00080354	0.00070859	0.00169885	0.00073413
2.588	0.01081023	0.00538665	0.00631249	0.00985006	0.00597615	0.01224038	0.00783254	0.00411341	0.00655103	0.00481883
2.647 (Methionine/Unknown)	0.0006251	0.00119454	0.00074254	0.00081768	0.00060214	0.00093448	0.00095236	0.00075102	0.00074162	0.00086258
2.662 (Succinic acid)	0.00059076	0.000315	0.00087586	0.00069187	0.00039573	0.00075123	0.0007029	0.00032153	0.00032032	0.00034134
2.742 (Sarcosine)	0.00017734	0.00015804	0.00011359	0.00028462	0.00019986	0.0001317	0.00021853	0.00009224	0.00021181	0.00033692
2.801	-0.00002694	-0.00000201	0.00007907	0.00002602	0.00003268	0.00013115	0.0000388	-0.0000122	-0.00006028	-0.00002977
2.831 (Citric acid)	0.00214856	0.00255951	0.00340321	0.00268489	0.001121	0.00379676	0.00340892	0.00139144	0.00184596	0.0037676
2.927 (Dimethylglycine)	0.0003523	0.00030975	0.00033552	0.00061121	0.00035708	0.00072788	0.00055218	0.00041	0.00066517	0.00065742
3.002 (Citric acid)	0.00965679	0.01058561	0.01047774	0.00991044	0.00486498	0.00854903	0.00866736	0.00550745	0.00521486	0.00755764
3.052 (Creatine)	0.00503777	0.00470919	0.00402342	0.00480677	0.0024064	0.00384909	0.00344436	0.003465	0.00329358	0.00543851
3.105	0.00050484	0.00072305	0.00082191	0.00070409	0.00047377	0.00067777	0.00083129	0.00062091	0.00050371	0.00049204
3.127 (Creatinine)	0.00210196	0.00147222	0.00128541	0.00158575	0.00158156	0.00151677	0.00179906	0.00155649	0.00096313	0.00225839
3.141 (DMSO2)	0.00073188	0.00096069	0.00074309	0.00112508	0.0006076	0.00078876	0.00141198	0.00100466	0.00047524	0.00115703
3.188 (Acetylcarnitine/Choline)	0.00721086	0.00448679	0.00393131	0.0097537	0.00152589	0.00577506	0.02080991	0.00328875	0.0096234	0.01402773
3.212 (Glucose)	0.0103889	0.01224055	0.01248105	0.01031994	0.00801365	0.00859913	0.01155127	0.01134751	0.01422299	0.01397615
3.221 (Carnitine)	0.00489235	0.0062378	0.00484756	0.00321699	0.00229066	0.00359282	0.0062694	0.00421823	0.00464837	0.00693646
3.239 (Glucose)	0.03680932	0.04300447	0.04360641	0.03938879	0.02247366	0.03172949	0.03764754	0.03978431	0.05066548	0.04768075
3.262 (Betaine/Myo-inositol)	0.00477437	0.0062451	0.00783742	0.01126345	0.00303952	0.00959516	0.00700006	0.00617772	0.01387562	0.00737595
3.283 (Myo-inositol)	0.00091857	0.00065998	0.00061842	0.00145656	0.00046242	0.00118657	0.00088315	0.00063697	0.00049662	0.00096743
3.333	0.00112292	0.00251627	0.00254233	0.0050517	0.00118268	0.00448138	0.00448445	0.00213678	0.00140835	0.00301012
3.350 (Methanol)	0.00898729	0.00882535	0.00700584	0.01246084	0.00394126	0.00859685	0.01570318	0.0062395	0.00688499	0.00914455
3.394 Glucose)	0.07107502	0.0845372	0.08781238	0.0781374	0.04750811	0.08925666	0.07729944	0.07760032	0.08971551	0.08001483
3.456 (Glucose)	0.07468796	0.08792884	0.09115201	0.07857687	0.05011819	0.06574422	0.07907833	0.07869519	0.0976583	0.08832215
3.518 (Glucose)	0.04176404	0.04141504	0.0433347	0.03996105	0.16510878	0.03445584	0.04779245	0.03710302	0.05094421	0.0461871
3.561	0.00532978	0.00269738	0.00407935	0.00594833	0.00458847	0.00571727	0.01518269	0.00261571	0.00590401	0.00415113
3.599 (Myo-inositol/Unknown)	0.00727562	0.00391683	0.00211898	0.00832132	0.0023495	0.00306635	0.00730324	0.0044484	0.00234237	0.00421296
3.647 (Ethanol)	0.01291919	0.00646248	0.0079406	0.0108026	0.00820912	0.02032948	0.02694931	0.00659417	0.0098027	0.00804731

3.728 (Glucose)	0.11547019	0.123695	0.12108326	0.11222848	0.07096731	0.08559927	0.1162154	0.11018029	0.13324107	0.12592136
3.783	0.00209598	0.00147115	0.00324633	0.00162221	0.00167503	0.0013505	0.00349281	0.00137543	0.00177993	0.00134841
3.801	0.00811525	0.00919615	0.00970584	0.00793693	0.00609026	0.00665186	0.00877759	0.00834763	0.01074976	0.01027887
3.830 (Glucose)	0.04112234	0.04917427	0.04478433	0.03952542	0.0274247	0.03229424	0.04098949	0.04175029	0.04761518	0.04252074
3.889 (Glucose)	0.04918996	0.0563821	0.05504318	0.05125649	0.28825383	0.03818203	0.04757974	0.04659403	0.05699622	0.05341372
3.921	0.00256268	0.00222053	0.00197443	0.00160139	0.00350556	0.00233723	0.0026288	0.0018904	0.00136183	0.00104139
3.951 (Betaine)	0.00179286	0.00210464	0.00215486	0.00208812	0.00058712	0.00184536	0.00270027	0.00213254	0.00315318	0.00197392
3.994	0.00627586	0.00728681	0.0085793	0.00893905	0.00239628	0.00735417	0.00678169	0.00595387	0.01060248	0.00823064
4.051 (Myo-inositol)	0.00277645	0.00127608	0.00097749	0.00174631	0.00150152	0.00044195	0.00310034	0.00138288	0.00177845	0.00280606
4.106 (Creatine)	0.00628246	0.0063432	0.00264991	0.00649762	0.00288218	0.00318133	0.00354023	0.00462288	0.00492909	0.0073543
4.144 (Proline/Unknown)	0.00036215	0.00015085	0.00336501	0.00094659	0.00018544	0.00069668	0.00086405	0.00000173	0.00019763	0.00013155
4.195	0.00147818	0.00209936	0.0008276	0.00159717	0.00135881	0.00104928	0.00504088	0.00171414	0.00123296	0.0011595
4.211	0.00035736	0.00063204	0.00113759	0.00113523	0.0003204	0.00184732	0.00123858	0.00053639	0.00069654	0.00020635
4.239	0.00028324	0.00028877	0.00168723	0.00300103	0.00057881	0.00827424	0.00011621	0.00045975	0.00192479	0.00075126
4.287 (Creatinine)	0.00109001	0.00059148	0.00064604	0.0007985	0.00084206	0.00036245	0.00044041	0.00076239	0.00046769	0.00143752
4.365 (Lactic acid)	0.04616405	0.02688862	0.03786939	0.03563477	0.03000913	0.03014947	0.02413517	0.06024023	0.02469841	0.02380855
4.512	-0.00002992	-0.00019488	-0.00006488	-0.00001763	-0.00009842	-0.0006224	-0.00020837	0.00004311	0.00011177	-0.00004923
4.632 (Glucose)	0.0167022	0.02281718	0.02206376	0.01610608	0.01070265	0.01507121	0.02067759	0.01960816	0.02417601	0.02433159
5.169 (Mannose)	0.00042247	0.00021941	0.00044215	0.00027862	0.00048049	0.00037188	0.00009794	0.00020289	0.00055896	0.00048543
5.221 (Glucose)	0.0243854	0.0282367	0.03007501	0.02539695	0.01577438	0.01844354	0.0239175	0.02495151	0.03708062	0.033335
6.091 (Inosine)	-0.00028895	-0.00024268	-0.0001944	-0.00015935	-0.00009981	-0.00037521	-0.00046734	-0.0001823	0.00000701	-0.00005184
6.890 (Tyrosine)	0.00184804	0.00292948	0.00185015	0.00204761	0.00081688	0.00143477	0.00140888	0.00182223	0.0017728	0.00294266
7.185 (Tyrosine)	0.00232126	0.00308204	0.00216724	0.00245909	0.00113538	0.00183696	0.00198705	0.00198649	0.00215371	0.00297548
7.322 (Phenylalanine)	0.00296087	0.00259755	0.00156915	0.00299985	0.00146685	0.00217687	0.0028872	0.00175857	0.00214523	0.00338769
7.407 (Phenylalanine/Histidine)	0.00507779	0.00531571	0.00347703	0.00599828	0.00226068	0.00323898	0.00469368	0.00375408	0.00337316	0.00569918
7.653	0.00551063	0.00167688	0.00048671	0.00078361	0.00022485	0.0003517	0.00003943	0.0004947	0.00054604	0.00039714
7.792	0.00492271	0.00124104	0.00034459	0.00053336	0.00011125	0.00037493	-0.00019802	0.0003005	0.00051877	0.00045266
7.931	-0.00013368	-0.00004925	0.00010646	0.00012837	0.00080922	-0.00013809	0.00003613	-0.00001272	0.00001916	-0.00007719
8.224 (Hypoxanthine/Inosine)	0.00002081	0.00023413	-0.00001231	0.00036226	0.00016839	0.0001952	0.00036801	0.00030359	0.0002996	0.00033902
8.243 (Formic acid)	0.00780453	0.00125776	0.0006402	0.00099025	0.0006246	0.00133505	0.00490563	0.00094578	0.00261634	0.00249437
8.418 (Hypoxanthine)	-0.00109171	-0.00047918	0.00073734	0.00013297	0.00120307	-0.00098499	-0.00153513	-0.00052143	0.00010933	0.0003456
8.654 (Histidine)	0.00109756	0.0014971	0.00115742	0.00160988	0.00053445	0.00076296	0.00109437	0.0007503	0.00131542	0.00198316

A1.2 Bin reduction

Assignment of potential metabolites to each bin, and the subsequent removal of outliers and elimination of bins with non-biological or unknown information, is shown in Figure A1-2.

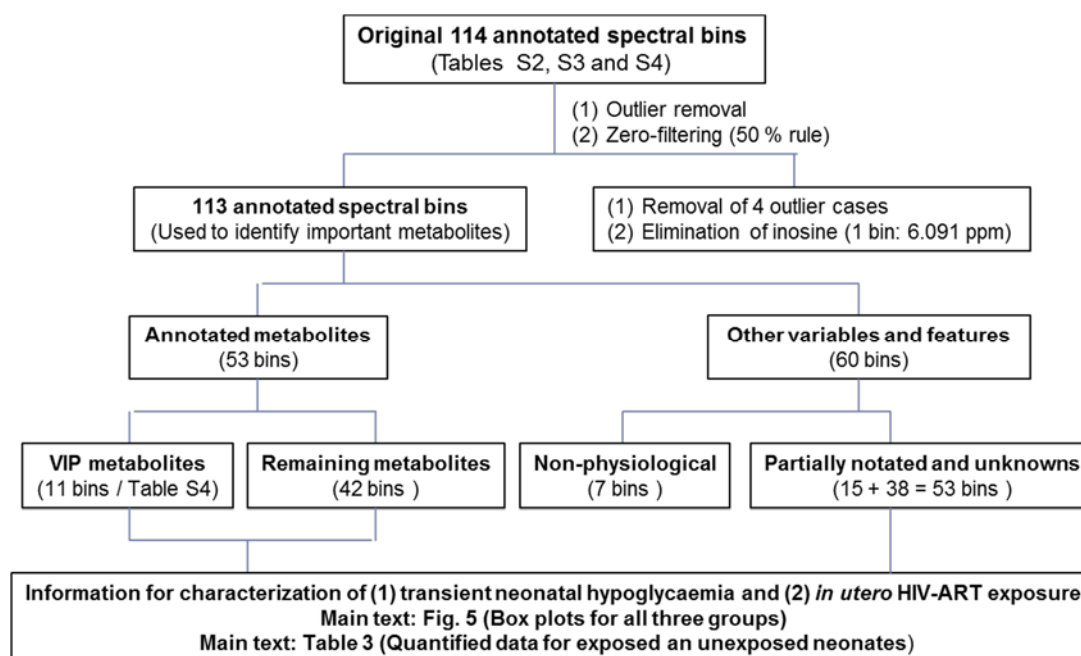


Figure A1–2: Flow diagram for the identification of variables from the spectral bins unequivocally assigned to metabolites

Only one variable — inosine (6.091 ppm) — was removed from the original data following zero reduction (50 % rule). Of the remaining 113 bins, 53 contained spectral information that could be assigned unequivocally to a single metabolite. These bins served to identify indicators of transient neonatal hypoglycaemia.

The remaining 60 bins contained 7 bins linked to 4 non-physiological variables: naturally occurring ^{13}C -lactic acid (2 bins); propylene glycol (2 bins), most likely a contaminant from the clinical procedures used to obtain cord blood; methanol (1 bin), formed in samples stored for some time at low temperatures; and ethanol (2 bins). The origin of ethanol from human metabolism is controversial (Logan & Jones 2000), although it has been observed in cord blood (Tea *et al.* 2012). The investigation was focused on variation within physiologically known endogenous metabolites. Thus, propylene glycol, methanol, ethanol and ^{13}C -lactic acid were excluded from further statistical analyses, although these four variables could be assigned unequivocally.

Finally, the remaining 53 bins, which contained 38 bins that could not be linked to any known metabolite, and 15 bins with only partial information on variable content (e.g., two alternative substances) were used to compare the unexposed and exposed neonatal groups. Neither the PCA (Figure A1-3a) and PLS-DA (Figure A1-3b) multivariate analyses nor the univariate Volcano plot (Figure A1-3c) indicated differentiation between these two neonatal groups. We conclude that the 53 bins do not contain information that indicates an effect of exposure to HIV-ART on the groups in combination (multivariate comparison) or as individual features (univariate analysis).

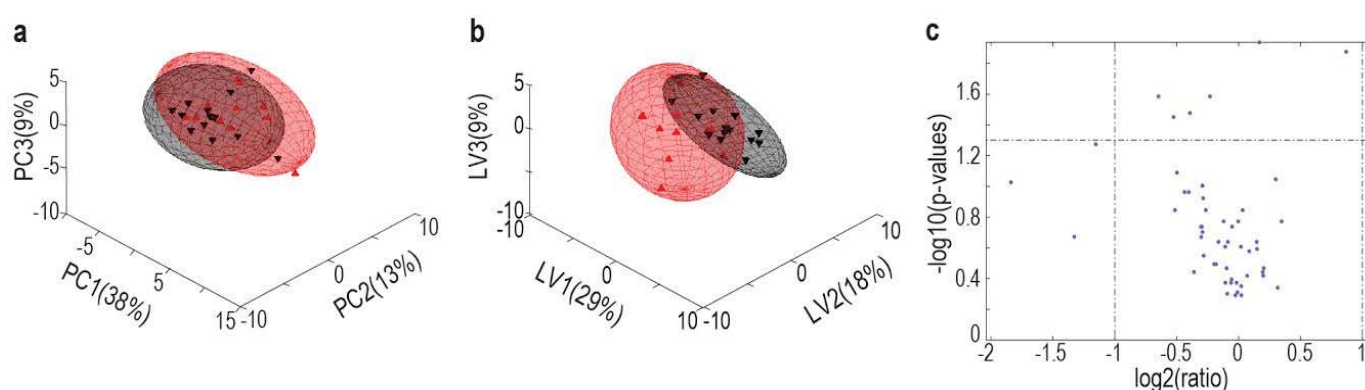


Figure A1-3: PCA (a), PLS-DA (b) and a Volcano plot (c), based on the data from the 53 bins for 15 unannotated and 38 unknown features, derived for the exposed (Table A1-2) and unexposed (Table A1-3) neonatal groups. Exposed groups are shown in red and the unexposed groups in black in (a) and (b) respectively.

A1.3 Outlier removal

A1.3.1 Procedure for identification of outliers

Following compilation of the spectral data (Tables A1-2 – A1-4), case reduction was applied to all three experimental groups. Four outliers were identified using a 95 % confidence region in a Hotelling's T^2 analysis (Figure A1-4a – A1-4c) in conjunction with the PCA score plots (Figure A1-4d–f) for the respective exposed, unexposed and infant groups. Supporting clinical and analytical justifications were used to remove cases. (1) Two cases in the infant group — one showed an abnormally high presence of suberic acid (Figure A1-5) and peaks for unknown substances, and the other had

several unidentified aberrations in several of the spectral bins. (2) One true outlier (by definition observed using both methods for outlier detection) in the unexposed neonatal group. (3) Two cases, one from the exposed neonatal group and the other from the infant group, were collected in vacutainers coated with EDTA, which caused some contamination (Figure A1-6). In addition, the presence of methanol and ethanol was observed in some samples (see A1.3.4), although they were not associated with outliers.

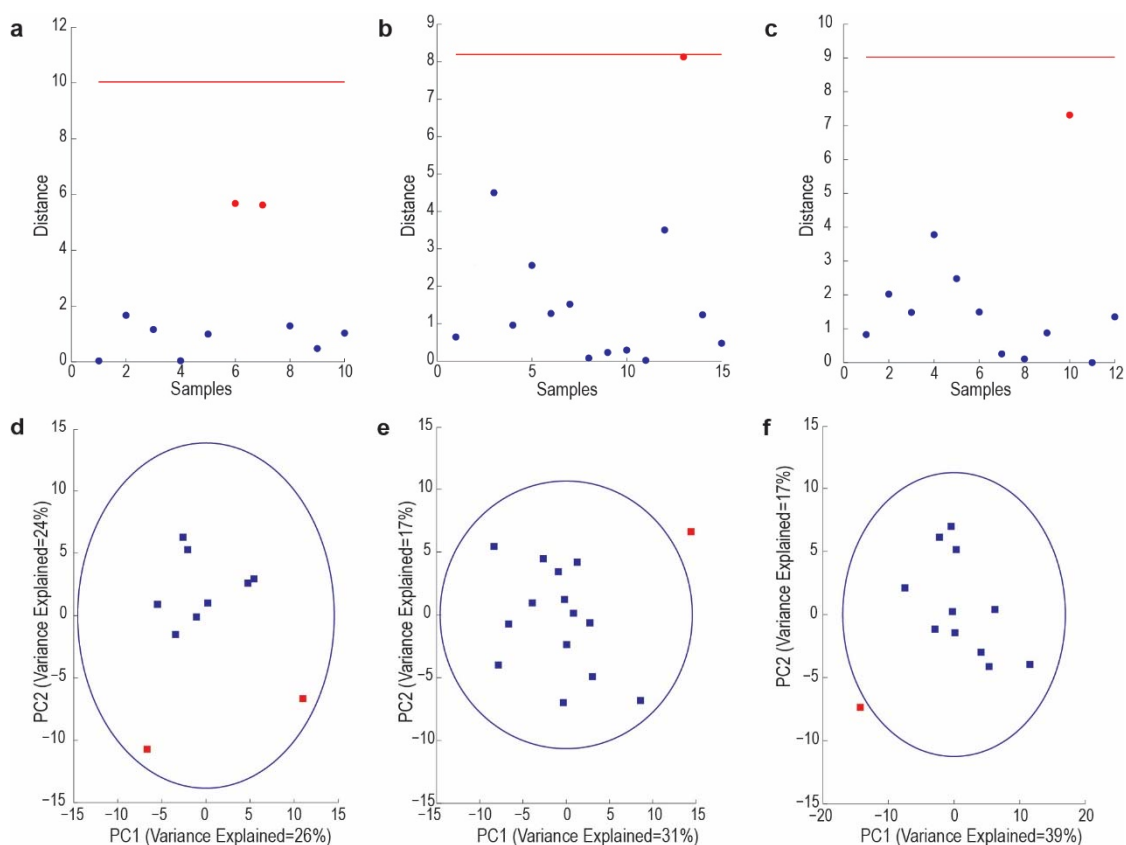


Figure A1-4: Outlier detection by Hotellings T^2 -test (**a**, **b** and **c**) and PCA (**d**, **e** and **f**). Data in **a** and **d** represent infant controls, in **b** and **e** the unexposed neonates, and in **c** and **f** neonates exposed to HIV-ART

A1.3.2 Single outlier in the Infant group.

The control group (n = 10) in this investigation were infants, whose plasma samples were sent to the Potchefstroom Laboratory for Inborn Errors of Metabolism (www.pliem.co.za) intended for follow-up analysis to confirm or reject suspicion of an inborn error of metabolism. After comprehensive MS metabolite analysis, their samples proved to be negative for any inborn errors of metabolism. Parental and ethical approval was obtained to use these samples in further studies. After NMR and statistical analysis were performed on these infant samples, a Hotelling's T^2 -test (Figure A1-4a) indicated two of the assigned infant bins to be outliers, which were identified to be suberic acid (Figure A1-5) and EDTA (Figure A1-S6). These cases were removed from further data analysis.

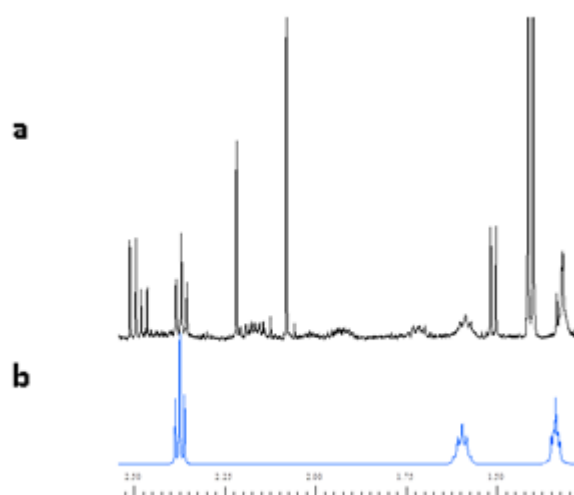


Figure A1-5: ^1H -NMR spectrum of suberic acid [2.37 ppm (triplet)], the only outlier detected in the infant group and in only one of its samples. The upper spectrum (**a**) is a section of the spectral image from the plasma sample of one of the infants; (**b**) represents the spectrum of the pure compound, from the Bruker library, used to confirm the presence of the metabolite thereby identified

A1.3.3 Outlier in the newborn group exposed to HIV-ART

Univariate statistical analysis, Hotellings T^2 -test, detected an outlier in the neonatal group exposed to HIV-ART. This outlier was identified as EDTA, a widely used blood anticoagulant in blood collection. The corresponding cord blood samples were

collected at birth in heparin vacutainers. The outlier detected is an indication of a contaminated sample, which was collected using a tube coated with EDTA (Figure A1-6). EDTA is known to interfere with the spectra in ^1H -NMR and can combine chemically with some of the metabolites present in the experimental sample, leading to the loss of biochemical information about metabolites of interest (Barton *et al.* 2010; Ross *et al.* 2007).

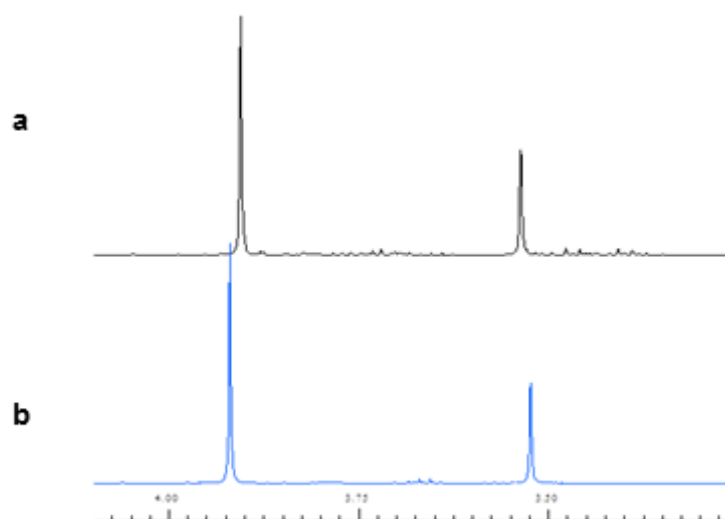


Figure A1–6: ^1H -NMR spectra of cord blood sample from a neonate exposed to HIV-ART, and of EDTA. In spectrum (a) the conspicuous exogenous substance in this cord blood sample was identified to be EDTA [3.90 ppm (singlet)]. The lower spectrum (b) is that of pure EDTA used to identify and to confirm the presence of EDTA

A1.3.4 Methanol and ethanol in cord blood samples

The presence of ethanol was substantiated by comparing the spectra of cord blood from both groups of neonates with that of an external ethanol standard (Figure A1-7a). The presence of ethanol in cord blood has been reported in both non-NMR (Logan & Jones 2000) and NMR (Tea *et al.* 2012) investigations, and interpreted as being of unknown or endogenous origin. Although ethanol is not used in standard operational procedures of clinical practice in the neonatal ward at the AMC, its incidental use in cleaning the umbilical cord cannot be excluded. We therefore excluded ethanol from our data matrix, justified as contamination. Methanol (Figure A1-7b) was excluded for similar reasons.

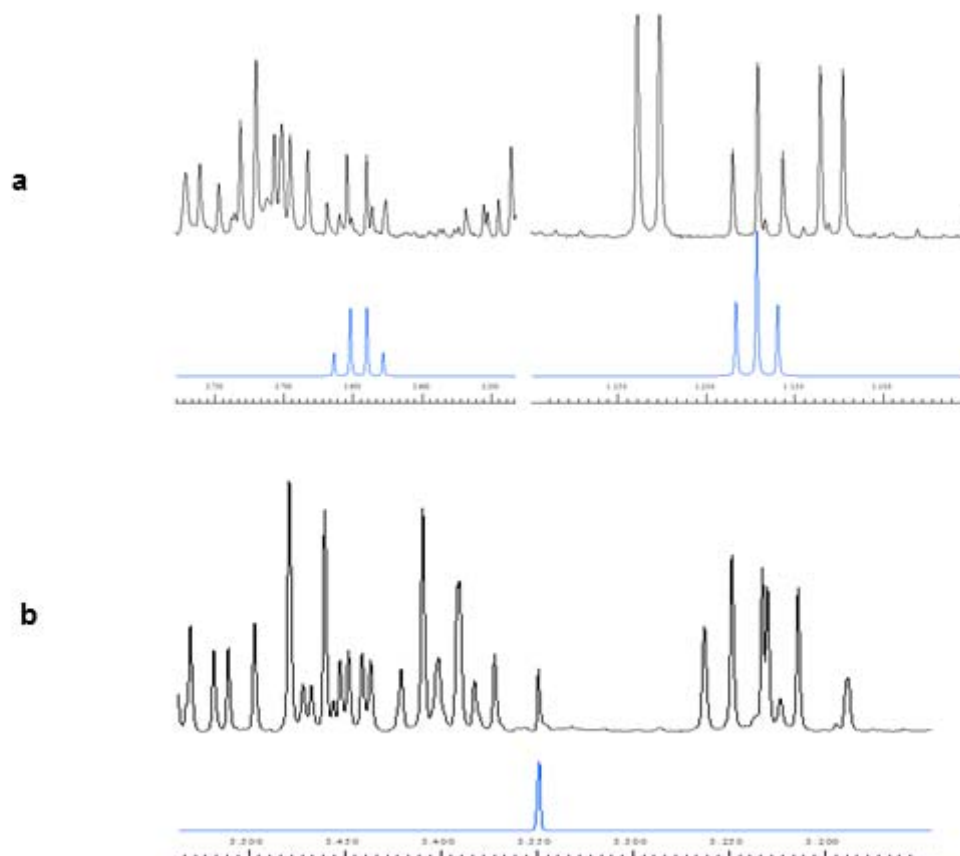


Figure A1–7: ¹H-NMR spectra of ethanol and methanol. These two alcohol metabolites were selected randomly from both groups of neonates and together were scaled according to an internal standard (IS) peak (TSP). Spectral windows of 1.15–1.19 ppm and 3.62–3.675 ppm were selected, representing the methyl [1.17 ppm (triplet)] and methylene [3.64 ppm (quartet)] groups, respectively, for ethanol. **(a)** As an example of these spectral windows, a cord blood sample (top panel) is compared with the spectrum of an ethanol standard (lower panel). **(b)** Similarly, a spectral window of 3.15–3.55 ppm was selected to represent the methyl group [3.35 ppm (singlet)] for methanol and compared with the spectrum of the methanol standard (lower spectra)

A1.3.5 Indications of medication

Propylene glycol (PG) is a commonly used drug excipient (a pharmacologically inactive substance). In medication it is used as a drug solubilizer and stabilizer in topical, oral and injectable medications. PG was present in all of our sample bin data (for both neonates and infants) of our investigation (Figure A1-8). Since our concern was about the physiological variation of known metabolites, and PG is an exogenous substance, it was removed from further statistical analysis of our bin data. In our cord blood samples its origin is due to intrauterine maternal-fetal transfer of medication and possibly diet, PG acts as a food additive and can cross the placenta. In infants its presence could be due to oral administration of paracetamol, a known drug often administered to infants (De Cock *et al.* 2012; Allegaert 2012) .

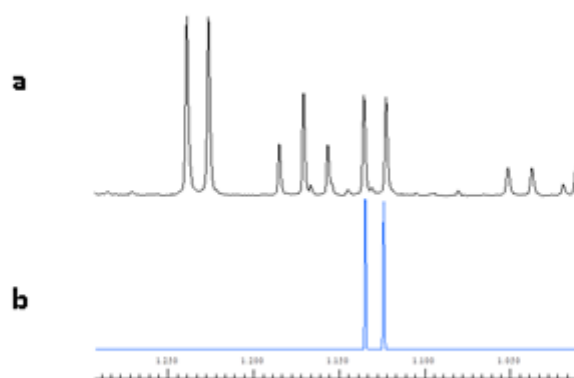


Figure A1–8: ^1H -NMR spectra of propylene glycol. The upper spectrum (**a**) is of a randomly selected sample from our investigation, to indicate the presence of propylene glycol [1.13 ppm (triplet)]. The lower spectrum (**b**) represents the propylene glycol standard used for identification and confirmation

Table A1–5: Eleven metabolites associated with spectral bins presenting with the most significant measures for discriminating between infant and unexposed experimental groups

Variable	Measures of significance between groups				Spectral descriptive statistics			
	Mann–Whitney test	Mann–Whitney test	Fold change	PLS-DA 2 VIP	Mean Infant	SD Infant	Mean Control	SD Control
Variable (ppm)	p-value	Effect size						
Carnitine (3.22)	0.0001	0.79	-2.46	3.34	0.0052	0.0012	0.0021	0.0006
Glucose (5.22)	0.0001	0.78	-2.67	2.45	0.0284	0.0048	0.0106	0.0076
Lactic acid (1.41)	0.0001	0.79	3.14	2.12	0.1083	0.0418	0.3402	0.1144
Phenylalanine (7.32)	0.0009	0.66	-1.73	2.24	0.0025	0.0006	0.0015	0.0004
Glutamine (2.47)	0.0012	0.65	-2.76	1.76	0.0101	0.0020	0.0037	0.0036
Formic acid (8.24)	0.0029	0.59	-3.68	1.66	0.0027	0.0025	0.0007	0.0003
Creatinine (4.29)	0.0009	0.66	1.85	1.61	0.0008	0.0003	0.0014	0.0004
Leucine (0.95)	0.0043	0.56	-1.40	1.44	0.0097	0.0023	0.0069	0.0016
Tyrosine (6.89)	0.0015	0.63	-1.49	1.35	0.0021	0.0006	0.0014	0.0003
Creatine (3.05)	0.0043	0.56	-1.37	1.25	0.0043	0.0008	0.0031	0.0006
*3-OH butyric acid (2.56)	0.0015	0.63	2.33	1.14	0.0011	0.0006	0.0025	0.0011

*The CH component of 3-OH-butyric acid (2.56 ppm AB system) was the bin which proved to be more significant in the univariate and multivariate statistics. However, because it stands alone, the methyl (CH₃) component of 3-OH butyric acid (1.23 ppm doublet) was quantified

SD, standard deviation

A1.4 Potential for ATP generation from catabolism of selective energy markers

Blood glucose concentrations as low as 2.5 mmol/L are common in healthy neonates at birth, and 1–2 hours after birth. These low concentrations are usually transient and asymptomatic phenomena, and are considered to be part of normal adaptation to postnatal life, following the disruption of the maternal-fetal metabolic association prior to birth. Most neonates compensate for this perturbed homeostasis due to transient hypoglycaemia by a process of allostasis. The allostasis response results in the production of ATP from progressively depleted glucose, supplemented by alternative fuels such as ketone bodies, which are released from fatty acid oxidation. A summary of the theoretical potential ATP generation by these processes, as applicable to the present investigation, is given in Table A1-6.

This table summarizes the molecules of ATP that can potentially be generated from the complete oxidation of glucose by glycolysis, the tricarboxylic acid cycle and oxidative phosphorylation (~38 molecules of ATP). Also indicated is the potential ATP production from lactic acid and 3-hydroxybutyric acid under conditions of glucose depletion. Acetone-based production of ATP is shown in the table but not used in the main text.

Table A1–6: Potential ATP generation from catabolism of selective energy markers

Pathway	Enzyme / Substrate	Oxidative mode of ATP generation	ATP molecules formed
1. Glycolysis			
	Glyceraldehyde-3-phosphate dehydrogenase	RC oxidation of 2 NADH	6 [malate shuttle]
	Phosphoglycerate kinase	Substrate level	2
	Pyruvate kinase	Substrate level	2
			Total 10
	Hexokinase and phosphofructokinase	ATP consumption	-2
			<u>Net from glycolysis 8</u>
2. Lactic acid / pyruvate shuttle			
	Pyruvate dehydrogenase [≡ lactic acid]	RC of 2 NADH	3
3. TCA and RC			
	Isocitrate dehydrogenase	RC of 2 NADH	3
	α-Ketoglutarate dehydrogenase	RC of 2 NADH	3
	Succinate thiokinase	Substrate level	1 [GTP]
	Succinate dehydrogenase	RC of 2 FADH ₂	2
	Malate dehydrogenase	RC of 2 NADH	3
			<u>Net from 2 and 3 15</u>
4. Ketone oxidation			
3-hydroxybutyrate	3-hydroxybutyrate dehydrogenase	RC of NADH	3
	Acetoacetic acid → 2 Acetyl-CoA	TCA and RC pathways	24
			<u>Net from 4 27</u>
Acetone	Acetoacetic acid → acetone + CO ₂	TCA and RC pathways	<u>Net from 4 12</u>

ATP potential = Glucose (μM x 38) + lactic acid (μM x 15) + 3-hydroxybutyric acid (μM x 27) [+ acetone (μM x 12)]

RC: Respiratory chain (oxidative phosphorylation); TCA: tricarboxylic acid cycle

A1.5.1. Integration of Methionine

Methionine, an amino acid in human metabolism that is indispensable in the diet, is potentially present in one of the bins which are responsible for the PLS–DA partial separation between HIV-ART exposed neonates and unexposed controls. In the NMR spectrum, methionine is evident as a singlet (2.13 ppm) and also as a triplet (2.67 ppm). In our investigation we were unable to quantify the singlet as it shared the same position as acetylcarnitine, a metabolite which too has a singlet at the 2.13 ppm position [it is referred to as 2.124 (methionine/acetylcarnitine) in Tables A1-2 – A1-4]. We therefore opted to quantify more open segment of the methionine spectrum, which is the triplet at 2.67 ppm. This proved to be easier said than done. As indicated in Figure A1-9, methionine is a widely spread triplet with a singlet corresponding to succinic acid (2.66 ppm) present within two of the methionine peaks, which excluded reliable quantification of this compound.

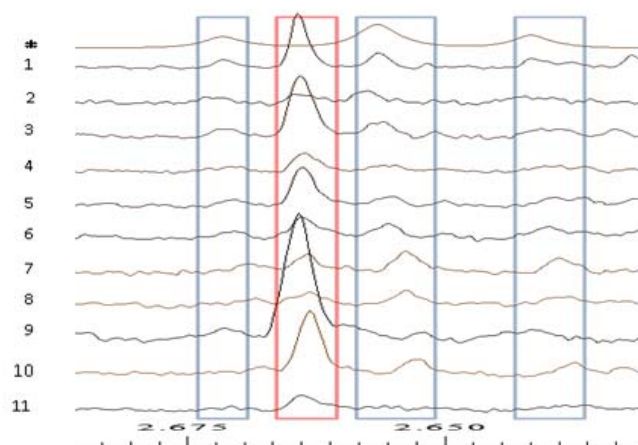


Figure A1–9: ^1H -NMR spectral integration of methionine and succinic acid. Methionine corresponds to a widely spread triplet at 2.67 ppm ($-\text{CH}_2-\text{S}$), indicated by the blue sections, present in all the 11 NMR specimens from neonates exposed to HIV-ART. Succinic acid is represented inside the red section centered at 2.66 ppm, a singlet that resides within the methionine triplet. Not all 11 spectra indicated the presence of methionine. For instance, in the 11th and 4th spectra there was no methionine to integrate. Methionine is therefore referred to as 2.647 (methionine/unknown) in Tables S2–S4. The spectrum marked with the asterisk (*) in the figure corresponds to the pure compound, which was used to identify methionine.

REFERENCES

- Allegaert, K. (2012). Propylene glycol in neonates; never prescribed, frequently administered, hardly evaluated. *Journal of Clinical Toxicology*, 2(9), 1–2.
- Barton, R. H., Waterman, D., Bonner, F. H., Holmes, E., Clarke, R., the PROCARDIS Consortium et al. (2010). The influence of EDTA and citrate anticoagulant addition to human plasma on information recovery from NMR-based metabolic profiling studies. *Molecular Biosystems*, 6(1), 215–224.
- Best, C. J. M., Gillespie, J. W., Yi, Y., Chandramouli, G. V. R., Perlmutter, M. A., Gathright, Y., et al. (2005) Molecular alterations in primary prostate cancer after androgen ablation therapy. *Clinical Cancer Research*, 11(19), 6823–6834.
- Cui, X., & Churchill, G. (2003). A. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology*, 4(4), 1–10.
- De Cock, R. F. W., Knillbabe, C. A. J., Kulo, A., De Hoon, J., Verbesselt, R., Danhof, M., & Allegaert, K. (2012). Developmental pharmacokinetics of propylene glycol in preterm and term neonates. *British Journal of Clinical Pharmacology*, 75(1), 162–171.
- De Meyer, T., Sinnaeve, D., Van Gasse, B., Tsiorkova, E., Rietzschel, E. R., De Buyzere, M. L., et al. (2008). NMR-based characterization of metabolic alterations in hypertension using an adaptive, intelligent binning algorithm. *Analytical Chemistry*, 80(10), 3783–3790.
- Engelke, U. F. H., Kremer, B., Kluijtmans, L. A. J., van der Graaf, M., Morava, E., Loeppky, F. J., et al. (2006). NMR spectroscopic studies on the late onset form of 3-methylglutaconic aciduria type I and other defects in leucine metabolism. *NMR in Biomedicine*, 19, 271–278.
- Engelke, U. F. H., Sass, J. O., Van Coster, R. N., Gerlo, E., Olbrich, H., Krywawych, S., et al. (2008). NMR spectroscopy of aminoacylase 1 deficiency, a novel inborn error of metabolism. *NMR in Biomedicine*, 21, 138–147.
- Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). Multi-and megavariable data analysis: Part I: basic principles and applications.

Logan, K. B., & Jones, A. W. (2000). Endogenous ethanol “auto-brewery syndrome” as a drunk-driving defense challenge. *Medicine, Science, and the Law*, 40(3), 206–215.

Ross, A., Schlotterbeck, G., Dieterle, F., & Senn, H. (2007). NMR spectroscopy techniques for application to metabonomics. In: Lindon J.C., Nicholson J. K., Holmes E., (ED.), *The Handbook of Metabolomics and Metabolomics* (pp. 55–122). Elsevier.

Tea, I., Le Gall, G., Küster, A., Guignard, N., Alexandre-Gouabau, M-C., Darmaun, D., & Robins, R. J. (2012). ¹H-NMR-based metabolic profiling of maternal and umbilical cord blood indicates altered materno-foetal nutrient exchange in preterm infants. *PLoS One*, 7(1), 1–12.

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, K. A., & van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142), 1–15.

Wevers, R. A., Engelke, U., Wendel, U., de Jong, J. G. N., Gabreëls, F. J. M., & Heerschap, A. (1995). Standardized method for high-resolution ¹H-NMR of cerebrospinal fluid. *Clinical Chemistry*, 41(5), 744–751.

KEMREP: A New Qualitative Method for the Assessment of an Analyst's Ability to Generate a Metabolomics Data Matrix by Gas Chromatography–Mass Spectrometry

Shayne Mason¹, Gontse P. Moutloatse¹, A. Marceline van Furth², Regan Solomons³, Mari van Reenen⁴, Carolus Reinecke^{1*} and Gerhard Koekemoer⁴

¹Centre for Human Metabonomics, Faculty of Natural Sciences, Private Bag X2006, North-West University (Potchefstroom Campus), South Africa; ²Department of Paediatric Infectious Diseases–Immunology and Rheumatology, Vrije Universiteit Medical Centre, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands; ³Department of Paediatrics and Child Health, Tygerberg Hospital, University of Stellenbosch, PO Box 19063, Tygerberg 7505, South Africa; ⁴Statistical Consultation Services, Faculty of Natural Sciences, Private Bag X2006, North-West University (Potchefstroom Campus), South Africa

Abstract: The analytical procedures required to generate a quantified metabolomics data matrix include many and widely different potential sources of error, complicating the generation of reliable data. The methods generally used to assess precision of such data all have distinct merits but some clear limitations as well. In this paper we describe KEMREP (kernel density estimation for the assessment of repeatability and reproducibility), a new method with the advantage and focus aimed specifically at analysis of the reliability of metabolomics data. Repeatability and reproducibility were assessed on gas chromatography-mass spectrometry (GC-MS) generated metabolomics data matrices produced by and between analysts and across laboratories, using cerebrospinal fluid (CSF) and urine as biological samples for analysis. KEMREP provides a visual overlay of the smoothed and scaled versions of the data from repeated samples for a direct and easy qualitative assessment of repeatability or reproducibility of a distinct chromatographic region (univariate) or for the experiment as a whole (multivariate). The KEMREP method can also be extended by the imposition of confidence bounds which provide lower and upper limits that indicate quantitatively whether the experiment was repeatable or reproducible at a predefined input coefficient of variation (CV). KEMREP is thus a novel approach which supplements existing methods of assessment of reliability of metabolomics data; provides a benchmark for assessing the quality of practical work performed by analysts; monitors the sequence of data pre-treatment steps; and tests the robustness of an experimentally designed protocol for metabolomics.

Keywords: Gas chromatography-mass spectrometry, kernel density, metabolomics, qualitative, repeatability, reproducibility.

INTRODUCTION

The terms “metabolome” [1] and “metabolomics” [2] have been well defined since they were first coined and evolved into an established field of research, which analytically culminates in producing a comprehensive qualitative and quantitative two-dimensional (a matrix) or three-dimensional (a tensor) distribution of all the metabolites present in a given biological system [3]. In this paper we focus on the qualitative aspect by presenting a novel way to analyse the reliability of metabolomics data by means of a qualitative assessment of kernel-density smoothed technical replicates (which we call KEMREP). Kernel density estimation is a non-parametric approach to estimate probability density through data smoothing. The main contribution of this approach is to assess the ability of analysts to generate a specified metabolomics matrix, examined by and between

analysts and across laboratories. This approach proved to be valuable to assess, in-house, the technical skills of researchers new to metabolomics analysis, and for periodical monitoring of such technical skills in analysts as a standard laboratory quality control practice. Finally, we propose how a quantitative conversion of the qualitative assessment using the KEMREP method can be introduced as well.

Analytical precision is of central importance in the generation of metabolomics matrices from biological samples [3, 4], capturing qualitative and quantitative information on compounds related to one or more complex biological perturbations under investigation. Samples may contain as many as hundreds of metabolites, which constitute the columns of the matrices. In most metabolomics investigations, these data points are designated as features, containing mainly true metabolites but also contaminants, artefacts and noise that should be eliminated in the process of data pretreatment [5, 6]. The quantified concentrations of metabolites are molecules of low molecular weight (<1500 Da) and products of cellular, metabolic and regulatory or pathological processes [4, 7, 8]. Their levels can be regarded as the response of biological systems to influences that are

*Address correspondence to this at the Centre for Human Metabonomics, School for Physical and Chemical Sciences (Division for Biochemistry), Faculty of Natural Sciences, Private Bag X6001, North-West University (Potchefstroom Campus), Potchefstroom 2531, South-Africa; Tel: 0027-18-299-2307; Fax: 0031-18-299-2316; E-mail: carools.reinecke@nwu.ac.za

logical systems to influences that are endogenous (e.g., genetic, homeostasis or age related) or exogenous (e.g., nutrition, infection or medication related). A metabolomics analysis can thus be described as a data-driven process to characterize the quantitative composition of these low molecular weight chemicals [2], with the objective of identifying those metabolites of interest in relation to the endogenous or exogenous perturbation under investigation. However, a need in current metabolomics is the further refinement and development of several aspects of metabolomics data generation and interpretation [9-11]. A key underlying aspect of all metabolomics investigations is the requirement that the analyst should produce data of a high quality [12], which is the central aspect addressed in this paper.

The analytical procedures required to generate a quantified metabolite profile include many and broadly different potential sources of error [13, 14], which complicates the generation and repeatability of metabolomics data; here, repeatability refers to the ability of a single analyst to produce consistent metabolite data, when using the same analytical method under identical experimental conditions. Apart from repeatability, reproducibility is another factor that affects the integrity of metabolomics matrices. Reproducibility here refers to assessment of repeatability of data generated analytically by more than a single independent source (i.e., evaluation of variation among laboratories, analysts, instruments and/or batches). Repeatability and reproducibility are thus two key terms used in assessing the overall variability of data and are essential in the validation of experimental results [12, 15], which requires the need to account for and ultimately eliminate or reduce unwanted analytical variability and focus on the perturbation of interest in order to produce statistical results that are biologically sensible [16, 17].

The relative standard deviation (RSD), or coefficient of variation (CV), are standard quantitative approaches used in metabolomics investigations that give insight into different sources of variation contributing to the total variation in the data and may be applied to monitor repeatability. The assessment of variability using CVs is, however, limited to single components within a measurement batch [18]; it also has the limitation of being unreliable at very low concentrations. Several univariate approaches [14, 19, 20] have been proposed to define measurement errors encountered in metabolomics experiments and to improve the assessment of the data; indeed, a low-level multivariate detection model [21] was advanced to assess the probability that certain peaks in gas chromatography-mass spectrometry (GC-MS) data are generated by the analytes, rather than by interferences. Van Batenburg *et al.* [13] recently presented a further development of the univariate Rocke-Lorenzato model [20]. It could be successfully applied to comprehensive metabolomics data, although it retained the limitation of describing the measurement error variance of single metabolites as a function of their responses.

Despite their limitations, all existing methods to assess precision of analytical data have distinctive merits in illuminating some specific aspect of the data under consideration. KEMREP has the advantage that it proposes an easily applicable tool for the convincing qualitative assessment of repeatability, as well as reproducibility. The design is based on

kernel density estimation and produces as output a smoothed chromatographic profile of the original data. Any GC-MS metabolomics data matrix generated from a single biological sample, subdivided into several identical aliquots for repeated analyses, may be used to evaluate the KEMREP method. The approach focuses on analytical variability, whereas the contribution of variability due to biological and/or environmental factors is nullified.

We selected a sample from two biofluids (urine from a healthy adult and cerebrospinal fluid (CSF) from a patient infected with, and treated for, tuberculous meningitis (TBM)) to illustrate the application of KEMREP. We used aliquots from the same sample to be analysed by the same or different analysts, using the same experimental protocol but generating the data on different GC-MS platforms, to address the issue of reproducibility. The GC-MS instruments chosen were almost identical, but used for different purposes in our laboratory: one was a multi-purpose platform, used primarily for general postgraduate student teaching and research purposes, whereas the other was dedicated to routine organic acid analysis for diagnosis of inherited metabolic disorders in our clinical-chemistry laboratory (www.pliem.co.za). The outline that we followed is presented in the flow diagram shown in (Fig. 1) and discussed below.

The practice of using a limited number of technical replicate samples to assess repeatability and reproducibility is a well-established and recognized approach in this kind of study (see Table 1). Such technical replicates are multiple samples, preprocessed by exactly the same protocol and used for consecutive injections into the GC-MS, to assess repeatability and/or reproducibility.

Achieving repeatability in metabolomics data is especially demanding [28] which reinforces the need to include qualitative assessments of repeatability and reproducibility at various stages of an experimental design in order to produce validated and reliable results, as addressed through the KEMREP approach presented here. Against this background, our protocol followed the typical experimental and theoretical procedures for organic acid profiling [8, 29, 30], presented in detail in the Supplementary Information (SI). This enabled us to succeed in devising an easy and practical approach for the qualitative assessment of the repeatability and reproducibility of the comprehensive analytical process used to generate a metabolomics matrix. We believe that this new method has the potential to be widely used in metabolomics practice, and will lead to its further development and refinement. To foster such initiatives, we include a detailed statistical description of the new method, as well as include as part of the SI: the original data, computer code for KEMREP and instructions on how to use the code, which is executable using the free statistical software program R (available for free download at <http://www.r-project.org>). This will enable fellow researchers to evaluate the appropriateness of the method for application on different metabolomics data sets.

MATERIALS AND METHODS

Experimental Design

The metabolomics pipeline, ranging from sampling to the point of new knowledge generated through the metabolomics

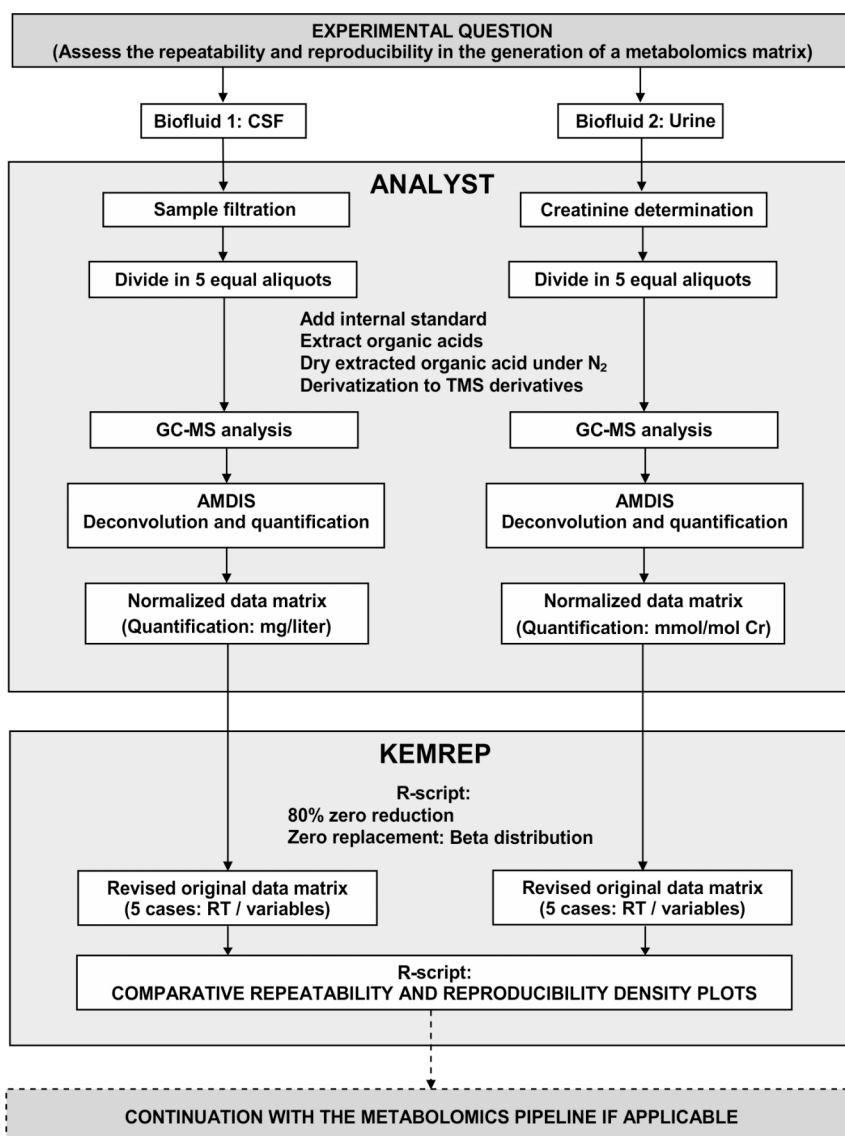


Fig. (1). Flow diagram illustrating experimental design focused on analytical procedures of the metabolomics pipeline followed in evaluating KEMREP.

Table 1. Selection of metabolomics-based studies that use technical replicates to assess repeatability and/or reproducibility.

	Technical Repeats (n)	Repeatability/Reproducibility Measure
Jiménez <i>et al.</i> [22]	5 – negative controls 4 – positive controls 4 – low concentration 4 – high concentration	RSD ratio in 11 qualitative chromatographic methods
Shepherd <i>et al.</i> [23]	4 – short term 9 – long term	CV
van Liempd <i>et al.</i> [24]	6 – single pooled 2 x 3 – double pooled	CV
Tredwell <i>et al.</i> [25]	3 x 6 different analytical methods (total = 18)	ANOVA, correlation analysis, pairwise rank correlations and CV
Tanaka <i>et al.</i> [26]	6	CV
t'Kindt <i>et al.</i> [27]	5 – pooled quality control	CV

approach, is one that has been well described not only from a holistic viewpoint [28–30], but also from specialized, focal aspects [12, 31–34]. Our design focuses on the analytical procedures in the pipeline, ranging from sample preparation to data processing, with particular attention being given to the analyst [35] (see flow diagram in Fig. (1)). By using the same sample for repeat analysis we eliminate the interference of biological variation, thereby making the initial sampling aspect of the pipeline uninfluential. Similarly, by proceeding to the point precluding the multivariate statistical analysis (data mining), we remove the need to quantify and apply biological interpretation to the data and focus on the analytical assessment, which is the prime focus of this paper.

Sample Selection: Background and Assumptions

Two distinctly different types of biofluid were used in this comparative investigation. The first sample was cerebrospinal fluid (CSF) obtained from a patient suffering from tuberculosis meningitis (TBM). The patient was a young African male (4 years and 10 months old at the time of sampling), who was admitted to Tygerberg Hospital, Department of Paediatrics & Child Health, University of Stellenbosch Medical School, South Africa, with clinical symptoms indicative of meningitis (persistent headache, cranial nerve palsy, sleepiness and strabismus) and assigned a Glasgow Coma Score of 13/15 [36]. The patient was treated with a broad-spectrum range of medication immediately upon admission after which biofluid samples were used for a definitive diagnosis. TBM stage II was confirmed by polymerase chain reaction (PCR) and microbiological culture analyses. At the time of sampling, direct ventricular drainage of CSF from the brain was performed to alleviate intra-cranial pressure, yielding approximately 20 ml of CSF. This sample, designated for analysis as well as assessment of repeatability, was divided into 1.0 ml aliquots, transported to our laboratory and stored at -80°C . After approximately 8 weeks in storage, five of the CSF aliquots were removed from -80°C storage and used for metabolomics analysis by one analyst, as described here.

The other sample used for the evaluation of KEMREP as an assessment tool was an early morning urine sample obtained from a single, healthy African adult female (27 years old) in an overnight fasted state under no form of medication. The urine sample can thus be considered as being uncomplicated by exogenous influences (i.e., a typical control sample).

GC-MS Analysis of Organic Acids

All matrices generated by the different analysts consisted of five aliquots of a single sample, whereas the number of variables differed from case to case. The variables in the original data set are considered features (components associated with peaks in the chromatogram defined as molecular entities with unique mass spectrometric m/z values at unique time points, or retention times, in a chromatogram) [37]. These compounds are later redefined as metabolites – chemical intermediates in the enzyme-catalyzed reactions of metabolism identified by the use of a database to link the retention time of a particular component to the accurate mass and structural formula derived from the model components

included in the database [8] or as defined substances – intact non-biological chemical components (e.g., derivatization by-products) ranging from those identified in the original chromatograms to some remaining in the final metabolomics data matrix, regarded as artefacts present in the original data [38].

In the case of the CSF analysis, each of the five aliquots were thawed and filtered using a 10,000 Da MWCO Centriscart-1 centrifugal unit by centrifugation for 15 min at 3000 rpm to remove any interfering proteins and bacteria or bacterial macromolecules present. Organic acids in the CSF were extracted using a standard operating procedure (SOP) as described previously [8]. The five prepared aliquots were analysed on two similar GC-MS instruments designated GC-MS-1 (used primarily for general postgraduate student teaching and research) and GC-MS-2 (used for dedicated routine organic acid analysis for metabolic diagnostic purposes) in two independent laboratories (see SI for detailed SOP and complete apparatus configuration). Both instruments had similar, but not identical, GC temperature oven programs, and so yielded different retention times and total run times. Similarly, the split ratios for each GC-MS apparatus differed marginally. These slight differences in configuration simulated a real-world situation in which two independent laboratories with the same type of instrument, measuring the same subset of the metabolome, have similar instrument settings but differ slightly in function. They also provided the opportunity to test the robustness of our proposed method of assessment of repeatability and reproducibility of our experimental data with respect to the SOP, GC-MS instruments and analyst. In the case of the urine analyses, the organic acids were extracted and derivatized by the same procedure as used for the CSF [8] and analysed on GC-MS-1. The Automated Mass Spectral Deconvolution and Identification System software program (AMDIS, version 2.66 from the National Institute for Standards and Technology) was used for spectral deconvolution and identification of features.

Prior to application of KEMREP, the data were normalized through quantification and logarithmic transformation. The final quantified concentrations of metabolites from the CSF are expressed as mg/liter and as mmol/mol creatinine for urine (Fig. 1). These are relative values by being volume-normalized (CSF) and creatinine-normalized (urine), respectively.

DEVELOPMENT OF THE KEMREP METHOD

Model Description

The KEMREP method aims to provide the user with a visual overlay of the information measured (i.e. smoothed chromatograms) from sample replicates in such a way that the user can quickly and easily assess repeatability or reproducibility at a distinct chromatographic region or for the experiment as a whole. This is accomplished by presenting a plot consisting of a smoothed and scaled version of the data with bounds that satisfy repeatability or reproducibility. We selected the data of the complex CSF sample used in our analyses and implemented some data reduction methods to assist in the explanation of the development of the statistical aspects of KEMREP.

The index i is used to represent a specific reproducibility item and the replicates of reproducibility item i will be denoted by j . Suppose there are I reproducibility items in total, then $n_1, n_2, \dots, n_I$ represent the number of samples (cases) for each reproducibility item. If p variables have been detected, then $t_{i,j,1}, t_{i,j,2}, \dots, t_{i,j,p}$ are the observed retention times for reproducibility item i , replicate j so that $t_{i,j,k} > t_{i,j,k-1}$ for $k \in \{2, 3, \dots, p\}$. The variable t is then used to represent a continuous retention time (i.e., t can assume any value between the smallest and greatest retention times observed in the data). Since the variables are measured on different value scales, and are all used simultaneously, some form of transformation is required to make the scales more comparable, yet retain the variation in which we are interested. For this purpose we use the shifted log transformation, with shift parameter set at one. The shifted log-scaled value is defined as $X_{i,j,t_{i,j,k}}$ for reproducibility item i , replicate j , at a fixed observed retention time $t_{i,j,k}$, where $k = 1, 2, \dots, p$.

The KEMREP method obtains the smoothed chromatogram for each replicate using an accumulation of the observed normalized and transformed values (the data used to contribute to illustrate the development of the KEMREP method is plotted in Fig. (2A)). Let $\theta_{i,j,k} = \sum_{q \in \{t_{i,j,1}, t_{i,j,2}, \dots, t_{i,j,k}\}} X_{i,j,q}$ for $k = 1, 2, \dots, p$, be the accumulated log-scaled values, which are those of each sample accumulated in order of increasing retention time. These accumulations are performed for each sample *individually*. An example of such an accumulation is presented in Fig. (2B). It should be noted that although this accumulation resembles a distribution function, it contains no frequency information. The only information contained within the “distribution” function is that of normalized transformed value accumulations as time elapses. Using linear interpolation (or monotone increasing spline interpolation – see Fritsch and Carlson [39]), we subse-

quently obtain the deterministic monotone increasing function $H(t)$ that can be used to map any retention time to an appropriate cumulative transformed value. Consider, as an example, a case of where the retention time versus the cumulative normalized and transformed values is plotted for five replicates of the same reproducibility item. If the experiment is repeatable, the cumulative values should all be similar when overlapped.

To find the smoothed chromatogram representation of the data (see Fig. 2C), we set up a fine, uniformly spaced grid (of size G) from the smallest normalized transformed value to the total accumulated transformed value for each sample. As an example, consider (Fig. 2B), where this grid will be constructed on the y-axis. We then use the inverse interpolation function $H^{-1}(\cdot)$ to find numerically the interpolated retention time values using the finely spaced grid as input (i.e., the inverse). This is illustrated visually (for one grid point) using arrows in Fig. (2B). The function $H(t)$ can therefore be considered as a model that can be used to generate data (retention times) given the functional form of the model. Note again that this function is *not* a probability model because the dependent variable contains no information that resembles frequencies (as it is constructed from a single replicate), but instead contains information concerning the intensity of the machine activity. The result of the interpolation is that we map the bivariate data (i.e., retention time and transformed value) to univariate data (i.e., retention time), in such a way that a higher density of retention times can be expected in areas where large values are observed, and a corresponding lower density in areas with less machine activity or where lower values were recorded. The weight allowed for each data point is thus a cumulative function. A histogram can now be constructed to visualize the retention time density that resulted from the function $H^{-1}(\cdot)$. Instead, we opt for a kernel density estimator which at a point t for data points $T_1, \dots, T_G$ is defined as:

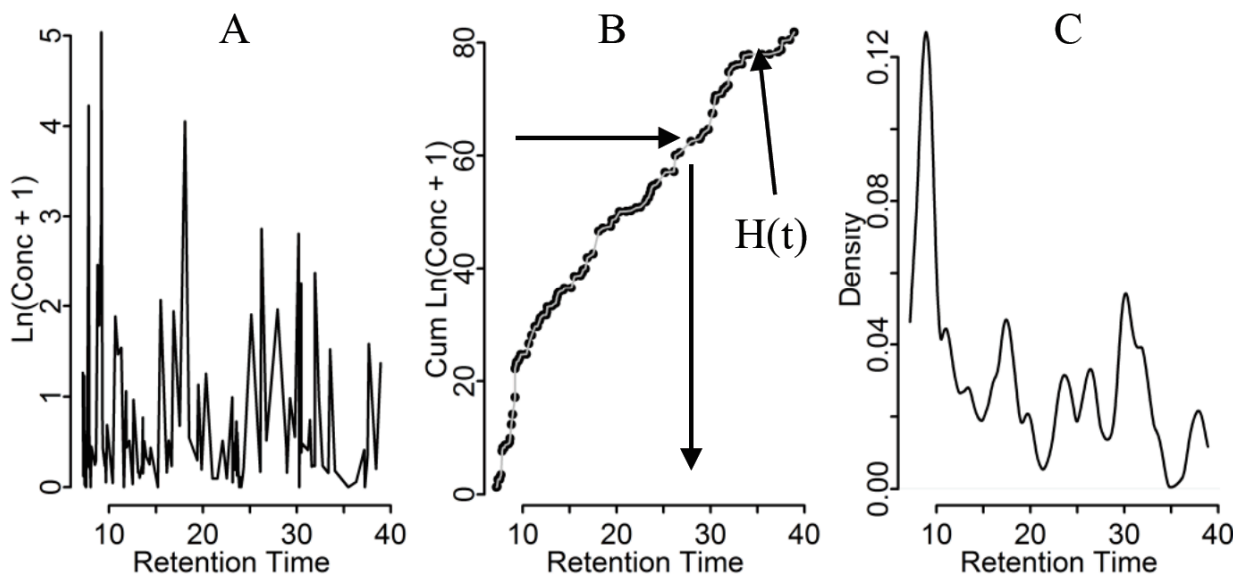


Fig. (2). A: Lineplot of retention time versus the normalized log-transformed data for a single sample (case). B: Accompanying scatterplot of retention times versus the cumulative normalized log-transformed data. The line $H(t)$ represents the linear (or monotone spline) interpolation function. The arrows represent an interpolated retention time. C: Density representation of the chromatogram (i.e. the KEMREP-smoothed chromatogram).

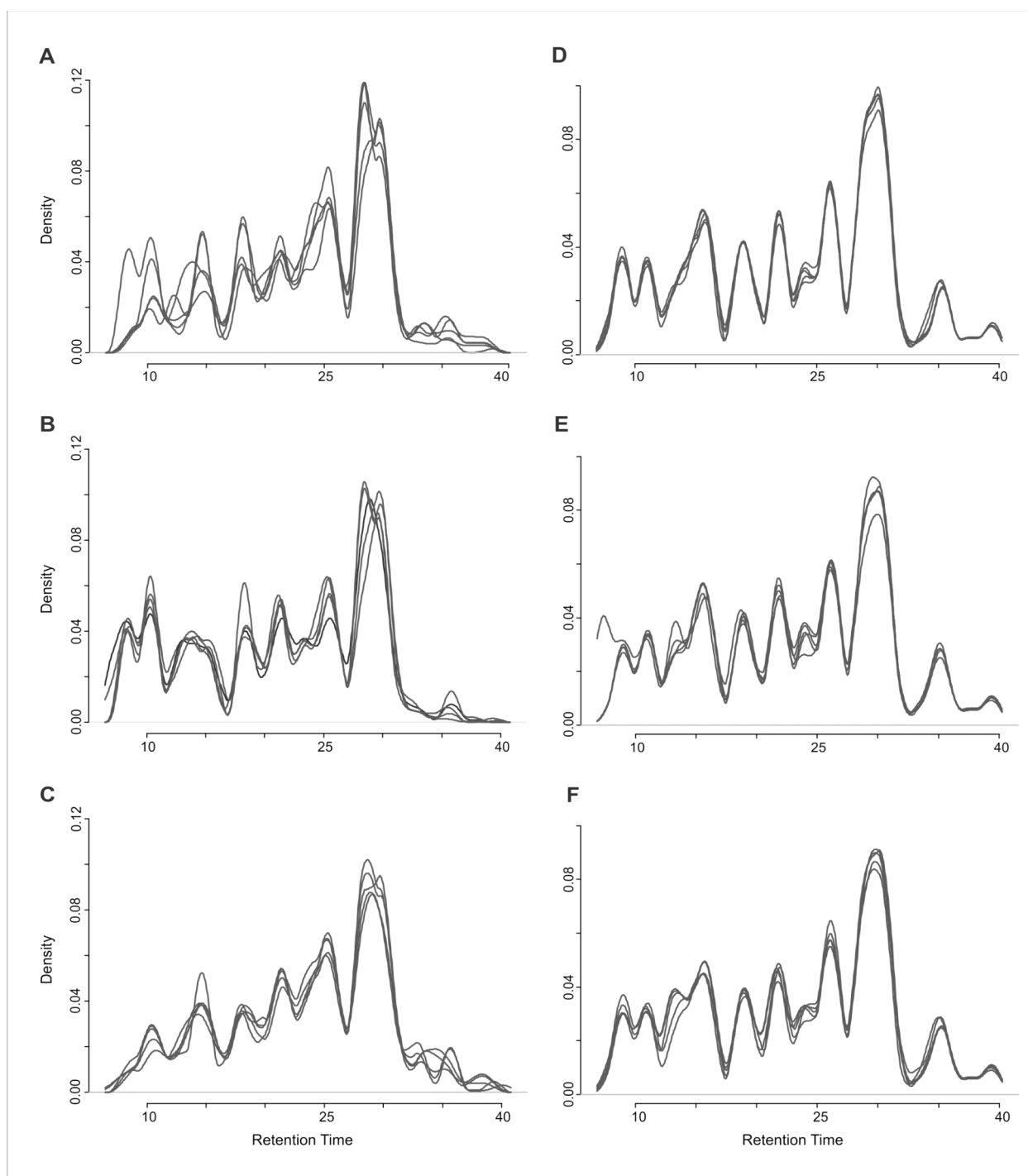


Fig. (6). Overlays of density plots of 5 repeat aliquots from a single urine sample. **A, B & C:** First attempt to generate the GS-MS data by analysts 1, 2 and 3, respectively. **D, E & F:** Second attempt to generate the GS-MS data by the same analysts, respectively. All analyses were conducted using the identical reagents, SOP and the same GC-MS apparatus.

biological sample, subdivided into several identical aliquots for repeated analyses, from any other biological sample or analytical method of analysis, may be used in the application of to evaluate the KEMREP method. To illustrate this statement we included the application of the KEMREP method to the determination of amino acids in 10 aliquots of a control urine sample and a freeze-dried, commercially available plasma sample spiked with amino acids, including a quantitative analysis (SI). The confi-

dence bounds, as well as the CV profiles of these analyses, give a further indication of the precision of the analyses and the topicality of the KEMREP method.

CONCLUSIONS

KEMREP, as developed and described here, has been shown to be a novel statistical method for an easy and direct qualitative assessment of the repeatability and reproducibility of metabolomics data matrices, based on its application to

real metabolomics data derived from different biofluids. The method can also be extended by the imposition of confidence bounds, which provides lower and upper limits that indicate quantitatively whether the experiment was repeatable or reproducible at a predefined input coefficient of variation (CV). Its application indicates several advantages in the evaluation of complex metabolomics data, emphasizing the potential wide application of KEMREP in metabolomics investigations. This paper thus highlights the ability of KEMREP in supplementing existing methods in assessing the reliability of metabolomics data generated by GC-MS. We are confident that the KEMREP method has the potential to be extended as: (1) a benchmark for determining analyst practical performance, (2) a means of monitoring experimental protocols and (3) evaluating the analytical robustness of metabolomics experimental design.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

S. Mason is a recipient of a Vrije Universiteit-NRF Desmond Tutu Doctoral Scholarship from the National Research Foundation (NRF) of South Africa. Funding for the project from the Technological Innovation Agency (TIA) of the South African Department of Science and Technology and from North-West University is likewise acknowledged.

REFERENCES

- Oliver, S. G.; Winson, M. K.; Kell, D. B.; Bagan, F. Systematic functional analysis of the yeast genome. *Trends Biotechnol.*, **1998**, *16*, 373-378.
- Fiehn, O. Metabolomics – the link between genotypes and phenotypes. *Plant Mol. Biol.*, **2002**, *48*, 155-171.
- Álvarez-Sánchez, B.; Priego-Capote, F.; Luque de Castro, M. D. Metabolomics analysis I: Selection of biological samples and practical aspects preceding sample preparation. *Trends Anal. Chem.*, **2010**, *29*(2), 111-119.
- Yin, P.; Peter, A.; Franken, H.; Zhao, X.; Neukamm, S. S.; Rosenbaum, L.; Lucio, M.; Zell, A.; Häring, H.-U.; Xu, G.; Lehman, R. Preanalytical aspects and sample quality assessment in metabolomics studies of human blood. *Clin. Chem.*, **2013**, *59*, 833-845.
- van der Greef, J.; Martin, S.; Juhasz, P.; Adourian, A.; Plasterer, T.; Verheij, E. R.; McBurney, R. N. The art and practice of systems biology in medicine: mapping patterns of relationships. *J. Proteome Res.*, **2007**, *6*, 1540-1559.
- Bruce, S. J.; Tavazzi, I.; Parlsod, V.; Rezzi, S.; Kochhar, S.; Guy, P. A. Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry. *Anal. Chem.*, **2009**, *81*, 3285-3296.
- Guy, P. A.; Tavazzi, I.; Bruce, S. J.; Ramadan, Z.; Kochhar, S. Global metabolic profiling analysis on human urine by UPLC-TOFMS: Issues and method validation in nutritional metabolomics. *J. Chromatogr. B*, **2008**, *871*, 253-260.
- Reinecke, C. J.; Koekemoer, G.; van der Westhuizen, F. H.; Louw, R.; Lindeke, J. Z.; Mienie, L. J.; Smuts, I. Metabolomics of urinary organic acids in respiratory chain deficiencies in children. *Metabolomics*, **2012**, *8*, 264-283.
- Zhang, A.; Sun, H.; Wang, P.; Han, Z.; Wang, X. Modern analytical techniques in metabolomics analysis. *Analyst*, **2012**, *137*, 293-300.
- Benton, H. P.; Want, E.; Keun, H. C.; Amberg, A.; Plumb, R. S.; Goldfain-Blanc, F.; Walther, B.; Reilly, M. D.; Lindon, J. C.; Holmes, E.; Nicholson, J. K.; Ebbels, T. M. D. Intra- and interlaboratory reproducibility of ultra performance liquid chromatography-time-of-flight mass spectrometry for urinary metabolomics data derived from different biofluids. *Anal. Chem.*, **2012**, *84*, 2424-2432.
- Worley, B.; Halouska, S.; Powers, R. Utilities for quantifying separation in PCA/PLS-DA scores plots. *Anal. Biochem.*, **2013**, *433*, 102-104.
- Ioannidis, J. P. A. Why most published research findings are false. *PLoS Medicine*, **2005**, *2*(8), 696-701.
- Van Batenburg, M. F.; Coulier, L.; Van Eeuwijk, F.; Smilde, A. K.; Westerhuis, J. A. New figures of merit for comprehensive functional genomics data: the metabolomics case. *Anal. Chem.*, **2011**, *83*, 3267-3274.
- Xu, F.; Zou, L.; Ong, C. N. Experiment-originated variations, and multi-peak and multi-origination phenomena in derivatization-based GC-MS metabolomics. *Trends Anal. Chem.*, **2010**, *29*(3), 269-280.
- Lay Jr., J. O.; Borgmann, S.; Liyanage, R.; Wilkins, C. L. Problems with the "omics". *Trends Anal. Chem.*, **2006**, *25*(11), 1046-1056.
- van den Berg, R. A.; Hoefsloot, H. C. J.; Westerhuis, J. A.; Smilde, A. K.; van der Werf, M. J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, **2006**, *7*, 142-157.
- Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Orešič, M. Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC Bioinformatics*, **2007**, *8*, 93-110.
- Urban, M.; Enot, D. P.; Dallmann, G.; Körner, L.; Forcher, V.; Enoh, P.; Koal, T.; Keller, M.; Deigner, H. P. Complexity and pitfalls of mass spectrometry-based targeted metabolomics in brain research. *Anal. Biochem.*, **2010**, *406*, 124-131.
- Smilde, A. K.; van der Werf, M. J.; Schaller, J. P.; Kistemaker, C. Characterizing the precision of mass-spectrometry-based metabolic profiling platforms. *Analyst*, **2009**, *134*, 2281-2285.
- Rocke, D. M. and Lorenzato, S. A two-component model for measurement error in analytical chemistry. *Technometrics*, **1995**, *37*, 176-184.
- Van der Voet, H.; de Boer, W. J.; de Ruig, W. G.; van Rhijn, J. A. Detection of residues using multivariate modelling of low-level GC-MS measurements. *J. Chemom.*, **1998**, *12*, 279-294.
- Jiménez, C.; Ventura, R.; Segura, J. Validation of qualitative chromatographic methods: strategy in antidoping control laboratories. *J. Chromatogr. B*, **2002**, *767*, 341-351.
- Shepherd, T.; Dobson, G.; Verrall, S. R.; Conner, S.; Griffiths, D. W.; McNicol, J. W.; Davies, H. V.; Stewart, D. Potato metabolomics by GC-MS: what are the limiting factors? *Metabolomics*, **2007**, *3*(4), 475-488.
- van Liempd, S.; Cabrera, D.; Mato, J. M.; Falcon-Perez, J. M. A fast method for the quantitation of key metabolites of the methionine pathway in liver tissue by high-resolution mass spectrometry and hydrophilic interaction ultra-performance liquid chromatography. *Anal. Bioanal. Chem.*, **2013**, *405*, 5301-5310.
- Tredwell, G. D.; Behrends, V.; Geier, F. M.; Liebecke, M.; Bundy, J. G. Between-person comparison of metabolite fitting for NMR-based quantitative metabolomics. *Anal. Chem.*, **2011**, *83*, 8683-8687.
- Tanaka, Y.; Higashi, T.; Rakwal, R.; Wakida, S.; Iwahashi, H. Quantitative analysis of sulfur-related metabolites during cadmium stress response in yeast by capillary electrophoresis-mass spectrometry. *J. Pharmaceut. Biomed.*, **2007**, *44*, 608-613.
- t'Kindt, R.; Morreel, K.; Deforce, D.; Boerjan, W.; Van Bocxlaer, J. Joint GC-MS and LC-MS platforms for comprehensive plant metabolomics: Repeatability and sample pre-treatment. *J. Chromatogr. B*, **2009**, *877*, 3752-3580.
- Fukusaki, E.; Kobayashi, A. Plant metabolomics: potential for practical operation. *J. Biosci. Bioeng.*, **2005**, *100*(4), 347-354.
- Goodacre, R.; Broadhurst, D.; Smilde, A. K.; Kristal, B. S.; Baker, J. D.; Beger, R.; Bessant, C.; Connor, S.; Capuani, G.; Craig, A.; Ebbels, T.; Kell, D. B.; Manetti, C.; Newton, J.; Paternostro, G.; Somorjai, R.; Sjöström, M.; Trygg, J.; Wulfert, F. Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics*, **2007**, *3*, 231-241.
- Brown, M.; Dunn, W. B.; Ellis, D. I.; Goodacre, R.; Handl, J.; Knowles, J. D.; O'Hagan, S.; Spasić, I.; Kell, D. B. A metabolome pipeline: from concept to data to knowledge. *Metabolomics*, **2005**, *1*(1), 39-50.
- Hendriks, M. M. W. B.; van Eeuwijk, F. A.; Jellema, R. H.; Westerhuis, J. A.; Reijmers, T. H.; Hoefsloot, H. C. J.; Smilde, A.

- K. Data-processing strategies for metabolomics studies. *Trends Anal. Chem.*, **2011**, 30(10), 1685-1698.
- [32] Liland, K. H. Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis. *Trends Anal. Chem.*, **2011**, 30(6), 827-841.
- [33] Katajamaa, M.; Orešič, M. Data processing for mass spectrometry-based metabolomics. *J. Chromatogr. A*, **2007**, 1158, 318-328.
- [34] Castillo, S.; Gopalacharyulu, P.; Yetukuri, L.; Orešič, M. Algorithms and tools for the preprocessing of LC-MS metabolomics data. *Chemometr. Intell. Lab.*, **2011**, 108, 23-32.
- [35] Milman, B. L. Identification of chemical compounds. *Trends Anal. Chem.*, **2005**, 24(6), 493-508.
- [36] Schutte, C. M.; van der Meyden, C. H. A prospective study of Glasgow Coma Scale (GCS), age, CSF-neutrophil, and CSF-protein and glucose levels as prognostic indicators in 100 adult patients with meningitis. *J. Infect.*, **1998**, 37, 112-115.
- [37] Tautenhahn, R.; Patti, G. J.; Kalisiak, E.; Miyamoto, T.; Schmidt, M.; Lo, F. Y.; McBee, J.; Baliga, N. S.; Siuzdak, G. metaXCMS: Second-order analysis of untargeted metabolomics data. *Anal. Chem.*, **2010**, 83, 696-700.
- [38] Wasim, M. and Brereton R. G. Determination of the number of significant components in liquid chromatography nuclear magnetic resonance spectroscopy. *Chemometr. Intell. Lab.*, **2004**, 72, 133-151.
- [39] Fritsch, F. N. and Carlson, R. E. Monotone piecewise cubic interpolation. *SIAM J. Num. Anal.*, **1980**, 17, 238-246.
- [40] Sheather, S. J. and Jones, M. C. A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Statist. Soc. B*, **1991**, 53(3), 683-690.
- [41] Wand, M. P. and Jones, M. C. *Kernel Smoothing*, Chapman and Hall, London, 1995.
- [42] Efron, B. and Tibshirani, R. J. *An Introduction to the Bootstrap (Monographs on Statistics and Applied Probability)*, Chapman and Hall, New York, 1993.
- [43] Yu, K.; Sheng, G.; Sheng, J.; Chen, Y.; Xu, W.; Liu, X.; Cao, H.; Qu, H.; Cheng, Y.; Li, L. A metabonomic investigation on the biochemical perturbation in liver failure patients caused by Hepatitis B virus. *J. Proteome Res.*, **2007**, 6, 2413-2419.
- [44] Mok, D. K. W.; Chau, F. T. Chemical information of Chinese medicines: a challenge to chemist. *Chemometr. Intell. Lab.*, **2006**, 82, 210-217.
- [45] Lu, H.; Dunn, W. B.; Shen, H.; Kell, D. B.; Liang, Y. Comparative evaluation of software for deconvolution of metabolomics data based on GC-TOF-MS. *Trends Anal. Chem.*, **2008**, 27(3), 215-227.
- [46] Bijlsma, S.; Bobeldijk, I.; Verheij, E. R.; Ramaker, R.; Kochnar, S.; Macdonald, I. A.; van Ommen, B.; Smilde, A.K. Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal. Chem.*, **2006**, 78, 567-574.
- [47] Williams, A.; Koekmoer, G.; Lindeque, Z.; Reinecke, C.; Meyer, D. Qualitative serum organic acid profiles of HIV-infected individuals not on antiretroviral treatment. *Metabolomics*, **2012**, 8, 804-818.
- [48] Meyer, M. R.; Weber, A. A.; Maurer, H. M. A validated GC-MS procedure for fast, simple, and cost-effective quantification of glycols and GHB in human plasma and their identification in urine and plasma for emergency toxicology. *Anal. Bioanal. Chem.*, **2011**, 400, 411-414.

Received: August 22, 2013

Revised: December 19, 2013

Accepted: January 13, 2014

ANNEXURE 3:

Supplementary Information

Metabolic risks of neonates at birth following in utero exposure to HIV-ART: The amino acid profile of cord blood

A3.1 Clinical information on HIV-ART exposed and unexposed neonates of cohort-1

Only biofluid samples (of serum derived from cord blood) with a volume of at least 1.0 mL (an analytical restriction) were selected to be used for NMR analysis (minimum 0.8 mL; Moutloatse *et al.* 2016), LC-QQQ [~50 µL; this study (cohort 1)] and UPLS-MS/MS [~1.5 mL; the complementary study (cohort 2)]. HIV-related parameters such as viral load, CD4 count, antiretroviral medication and demographic data were extracted by the same analyst (GPM) under the supervision of Dr Madeleine Bunders from records at the Academisch Medisch Centrum (AMC), the university hospital affiliated to the University of Amsterdam, in the Netherlands.

Table A3-1 summarizes the clinical parameters for the mothers of the ART-exposed neonates ($n = 12$), hereafter referred to as the exposed group. The mean age of these HIV-infected women was 32, ranging from 21 to 40 years at the time of delivery. There was a mean nadir CD4 cell count of 235 (range 40–920), and as part of the clinical follow-up during the pregnancy, CD4 T-cell count and plasma HIV-RNA were measured during the first, second and third trimesters of pregnancy. Almost all women had been diagnosed with HIV before this pregnancy and received combination-ART (cART) from the first trimester. The cART of the mothers was not identical, but reflected the real-world situation of ART adhered to by HIV-positive mothers according to prescription by their clinicians. According to European guidelines, the cART profile consisted of a nucleoside reverse transcriptase inhibitor (NRTI) backbone combined with either a protease inhibitor or the non-nucleoside reverse transcriptase inhibitor (NNRTI), nevirapine. The mean gestational age was 39.1 (38.3–41.1) weeks. Modes of delivery were vaginally ($n = 7$), vaginally with induction ($n = 1$), and caesarean section ($n = 4$).

Table A3-2 summarizes the clinical parameters of the exposed neonates (cohort 1). The mean birth weight of the newborns was 3449 (2900–4620) g. Inclusion criteria for the case study group was a gestational age of ≥ 35 weeks and birth weight [2500 g (appropriate for gestational age)]. The exclusion criteria were HIV-exposed neonates who tested positive to any other bacterial or viral infections detected. Venous umbilical cord blood samples were obtained within 10 min of delivery. Further detail on concentrations of individual serum metabolites, including glucose, lactic acid, 3-hydroxybutyric acid and hypoxanthine is reported in Moutloatse *et al.* 2016.

The controls consisted of 15 full-term singleton neonates, hereafter referred to as the unexposed cases, born vaginally after an uneventful pregnancy. The mode of delivery was comparable but not identical to the HIV-ART-exposed newborns. An inclusion criterion for this control group was healthy singleton babies with a gestational age of ≥ 37 weeks with a weight appropriate for gestational age and an uneventful pregnancy and delivery. Exclusion criteria for the control group were neonatal low birth weight, perinatal asphyxia, HIV-infected mothers and any clinical or genetic abnormality in the neonates.

Ethical approval for the study and for the use of the samples from the two neonatal groups was obtained from the Ethics Committee of the University of Amsterdam (see Moutloatse *et al.* 2016).

Table A3–1: Clinical information of the HIV-positive mothers for cohort-1, receiving ART during pregnancy

Age (years)	Ethnicity	Viral load ^a				CD4 count (10E ⁶ cells/L)			ART treatment regime (Y: Yes; N: No)					
		1 st	2nd	3 rd	Nadir ^b	1st	2nd	3rd	Combivir (Zidovudine + Lamivudine)	Kaletra (Lopinavir + Ritonavir)	Lamivudine	Nelfinavir	Nevirapine	Tenofovir
39	Caucasian	0	0	0	180	420	380	380	N	Y	N	N	Y	N
34	Black	0	0	0	200	540	460	500	Y	N	N	Y	N	N
21	Black	0	0	0	130	290	260	280	Y	N	N	N	Y	N
24	Black	0	0	0	160	600	680	580	Y	Y	N	Y	N	N
39	Black	0	0	0	330	490	510	660	N	N	Y	N	Y	N
24	Black	0	0	523	800	n/a	800	750	Y	Y	N	N	N	N
40	Black	0	0	0	n/a	790	720	850	N	N	Y	N	Y	N
29	Black	0	0	0	180	610	430	550	Y	N	Y	N	Y	Y
39	Black	28365	16210	0	180	180	230	200	Y	Y	N	N	N	N
30	Black	0	0	0	40	460	400	510	Y	N	N	N	Y	N
33	Black	0	0	0	170	440	430	410	Y	Y	N	N	N	Y
34	Caucasian	0	0	0	220	1300	970	950	Y	N	N	N	Y	N

^a Viral load: 0 = zero or undetectable

^b Nadir is lowest CD4 count

n/a, not applicable

Table A3–2: Clinical data details of the neonates *in utero* exposed to HIV-ART for cohort-1

Gender (M:male; F:female)	Ethnicity	Gestational age (weeks)	Mode of delivery*	Birth weight (g)	Apgar (1,5,10 min)	score	
M	Caucasian	39.4	Vaginal (Normal)	4620	8	10	
M	Black	38.5	Caesarean (Normal)	4090	9	10	
M	Black	40.2	Vaginal (Normal)	3170	8	10	
M	Black	39.2	Vaginal (Normal)	3560	6	8	10
M	Black	38.5	Caesarean (Normal)	3440	7	9	10
F	Black	38.3	Vaginal (Normal)	3035	10		
M	Black	40.1	Caesarean (Normal)	3465	10	10	
F	Black	39.3	Vaginal (Normal)	2900	9	10	
F	Black	38.3	Vaginal (Normal)	3185	9	10	
M	Black	37.6	Vaginal (Induced)	3850	10	10	
F	Black	39.1	Vaginal (Normal)	2565	4	7	8
M	Caucasian	41.1	Caesarean (Normal)	3506	8	10	10

*: Normal indicated mothers who gave birth without prior induction.

A3.2 Clinical information on HIV-ART exposed and unexposed neonates of Cohort-2

For cohort 2, 12 HIV-1-infected and 15 HIV-1-negative pregnant women were included. The baseline maternal group demographics (Table S3) were similar, except for ethnicity. In HIV-1-infected women triglyceride (1.8 mmol/L) and total cholesterol (5 mmol/L) levels were within upper-normal ranges for pregnant women; none of the HIV-1-infected women or in the control group were diagnosed with gestational diabetes during pregnancy. The HIV-1-infected women did not suffer from severe HIV-1-related immune deficiency, with a CD4 T-cell count in the first trimester exceeding 200×10^6 cell/L (Table A3-3). As expected after initiating cART, the CD4 T-cell counts further increased and viral load decreased; at delivery, viral load was undetectable in all women (Table A3-3). All mothers received cART in accordance with European AIDS Clinical Society guidelines including at least two nucleotide reverse transcriptase inhibitors (NRTIs) in combination with either a protease inhibitor (PI) or non-nucleotide reverse transcriptase inhibitor (NNRTI), with the majority of women receiving zidovudine, lamivudine and ritonavir-boosted lopinavir (Table A3-3).

None of the HIV-1-infected mothers transmitted the virus to their infant, which was verified by follow-up testing of HIV viral load in plasma. The HIV-1 exposed (HEU) and

HIV-1-unexposed (HUU) infants were similar with regard to gestational age, Apgar scores, mode of delivery and gender. As expected, birth weight was lower in the HEU infants, although this did not reach significance.

Table A3–3 Clinical information of the HIV-positive mothers receiving ART during pregnancy and for the HIV-ART neonates of cohort-2

Table 1: Patient Demographics

Clinical variables		Control mother (n = 15) & HUU-infants (n = 15)	HIV positive mother (n = 12) & HEU-infants (n = 12)	p-value
		Median (75% CI)	Median (75% CI)	
Mothers				
Age (years)		32.6 (30.7 - 33.2)	28.5 (28.2 -32.4)	0.50 ^a
Pre-pregnancy weight (kg)		64.5* (62.7 - 66.9)	68.5 (65.7 - 74.2)	0.41 ^a
Pre-pregnancy BMI		22.04* (22.3 - 24.1)	23.67 (22.6 - 26.1)	0.79 ^a
Total Triglycerides [#] (mmol/L)		-	1.80** (1.5 - 2.1)	
Total Cholesterol [#] (mmol/L)		-	5 ** (4.6 - 5.4)	
Glucose [#] (mmol/L)		-	4.30** (4.2 - 4.7)	
Ethnicity (W/B/SA)		9/4/2	3/9/0	0.03 ^b
Gestational Diabetes (Yes/No)		0/15	0/12	
HIV viral load (copies/mL)	trimester 1	-	528.5 (1269 - 9084)	
	trimester 2	-	50 (47.81 - 281)	
	trimester 3	-	0	
CD4 count (10E ⁶ cells/L)	trimester 1	-	455 (465.5 - 594.5)	
	trimester 2	-	520 (511.4 - 650.6)	
	trimester 3	-	700 (616.5 - 792.5)	
Infants				
Gestation period (weeks)		40 (37.63 - 39.97)	39.25 (38.74 - 39.82)	0.57 ^a
Birth weight (g)		3380 (3128 - 3494)	3085 (3009 - 3472)	0.34 ^a
APGAR score (5 min)		10 (9.5 - 9.9)	10 (9.1 - 9.9)	0.44 ^a
Mode of delivery (Vag/CS)		14/1	8/4	0.14 ^b
Gender (Male / Female)		8/7	7/5	0.55 ^b

^a Mann-Whitney U-test; ^b Fisher's exact test; [#] Measured at the end of the 2nd trimester; * (**) 1(2) missing value(s)

W - White; B - Black; SA - South Asian; Vag - Vaginal; CS - Caesarean section; CI – Confidence interval

A3.3 Acylcarnitine isotope solution preparation

Table A3-4: Acylcarnitine concentrations

ISOTOPE SOLUTION PREPARATION AND USAGE											
VERSION	N12.05										
[10,10,10-d3] Decanoyl-L-carnitine			5.3mg in 100mL dd H2O								
[12,12,12-d3] Dodecanoyl-L-carnitin			4.6mg in 100 mL ddH2O								
[14,14,14-d3] Tetradecanoyl-L-carni			4.9mg in 100 mL ddH2O								
[18,18,18-d3] Octadecanoyl-L-Carni			5.6mg in 100 mL ddH2O								
Date: 6 April 2009 Used all the new stock solutions of isotope prepared											
Used all the new stock solutions of isotope prepared (10 June 2008)											
ACYL CARNITINE ISOTOPE PREPARATION											
	Weighed mg	Mr Includes HCl	Mass of isotope	Mass of ester +H	Preferred loss	Vol mL	Stock Sol mmol/L nmol/μL	Dilutions: Take μL to 1000mL	Take μL to 200mL	Final conc(FC) μmol/L of nmol/m	Conc after Prep (x250μL/100μL)
Free carnitine (AC221) (Stock 2 Bottle 2) [methyl-d3]-L-carnitine.HCl, > 95%, > 98% d3 C7D3H12NO3.HCl MW 200.7	6.179	200.7	221.2	218.18	85	100	0.307872446	500	100	0.1539	0.3848
Acetylcarnitine (AC263) (Stock 2 Bottle 2) [d3]acetyl-L-carnitine.HCl, >95%, > 98% d3 C9D3H14NO4.HCl MW 242.7	5.011	242.7	263.21	260.19	85	100	0.206468892	500	100	0.1032	0.2581
Propionylcarnitine (AC277) [3,3,3-d3]propionyl-L-carnitine.HCl, > 95%, > 98% d3 C10D3H16NO4.HCl MW 256.7	5.476	256.7	277.22	274.2	85	100	0.213322945	500	100	0.1067	0.2667
isovalerylcarnitine (AC311) [d9]isovaleryl-L-carnitine.HCl, > 95%, > 98% d3 C12D9H14NO4.HCl MW 290.8	5.477	290.8	311.2	302.23	85	100	0.188342503	500	100	0.0942	0.2354
Octanoylcarnitine (AC347) [8,8,8-d3]octanoyl-L-carnitine.HCl, > 95%, > 98% d3 C15D3H26NO4.HCl MW 326.9	5.51	326.9	347.2	344.28	85	100	0.168553074	500	100	0.0843	0.2107
Palmitoylcarnitine (AC459) [16,16,16-d3]hexadecanoyl-L-carnitine.HCl, > 95%, > 98% d3 C23D3H42NO4.HCl MW 439.1	5.063	439.1	459.4	456.4	85	100	0.115304031	500	100	0.0577	0.1441
[10,10,10-d3] Decanoyl-L-carnitine	5.300	354.9	375.31	372.31	85	100	0.149337842		100	0.0747	0.1867
[12,12,12-d3] Dodecanoyl-L-carnitine	4.600	383	403.34	400.34	85	100	0.120104439		100	0.0601	0.1501
[14,14,14-d3] Tetradecanoyl-L-carnitine	4.900	411	431.37	428.37	85	100	0.119221411		100	0.0596	0.1490
[18,18,18-d3] Octadecanoyl-L-carnitine	5.600	467.1	487.44	484.44	85	100	0.119888675		100	0.0599	0.1499
Propionylcarnitine (AC277), Isovalerylcarnitine (AC311) and Octanoylcarnitine (AC347) are a mix - only 100ul of the mix											
VOLUMES USED FOR SAMPLE PREPARATION FOR CARNITINE ANALYSIS											
FOR URINE											
Sample:	10μL										
Isotope:	400μL of FC										
FOR SERUM/PLASMA											
Sample:	10μL										
Isotope:	400 μL FC										
FOR BLOODCARD											
Punch 1 x 1.495 μL Sample: 1.495 μL (1 x 3.175 mm Punch)											
Isotope: 250 μL of FC											
PICKERING STANDARDS											
The Pickering Stock Standard solution contains 0.250 nmol/μL of each amino acid.											
PREPARATION OF WORKING SOLUTION											
The following STANDARDS are prepared in Triplicate for each working batch:											
1. Make a 2:10 dilution with methanol.											
Example: Take 100μL ethanol and 25 μL Pickering.											
2. Take 30 μL of working solution and add 250 μL of FC isotope.											
3. For calculation purposes 6 μL is used of initial stock (0.250 nmol/μL*)											
Mr weight according to Herman Ten Brink document											
Isotope Mr weight includes HCl											
Please notice that molecular weights include the HCl (36.5)											

A3.4 Amino acid isotope solution preparation

Table A3-5: Amino acid isotope concentrations

ISOTOPE SOLUTION PREPARATION AND USAGE									
VERSION	N12.05								
New Citrulline	26.2.2013	10.05mg in 100mL metanol							
AMINO ACID ISOTOPE PREPARATION									
Amino Acids	Weighed	Mr		Vol	Stock Sol	Dilutions:		Final conc (FC)	Conc after Prep
	mg		Mass of Butyl Ester	ml	mmol/L	Take mL to 1000mL	Take µL to 200mL	µmol/l of nmol/ML	(x250µL/100µL)
					nmol/µl				
GLY	41.1	77.08	132.1	200	2.6660612	1.500	300	3.9991	9.9977
VALD8	40.08	125.15	174.15	200	1.6012785	2.500	500	4.0032	10.0080
ISOLEU	20.52	141.17	188.16	200	0.7267833	11.000	2200	7.9946	19.9865
MET	12.1	152.2	206.12	200	0.3975033	7.500	1500	2.9813	7.4532
PHE	40.63	170.19	222.15	200	1.1936659	2.000	400	2.3873	5.9683
GLU	39.9	152.13	260.19	200	1.3113784	2.500	500	3.2784	8.1961
TRP	20.1	209.23	261.16	100	0.4803326	3.000	600	2.8820	7.2050
LYS	41.4	150.19	203.18	200	1.3782542	5.000	1000	6.8913	17.2282
ARG	21.48	178.2	231.18	200	0.6026936	7.000	1400	4.2189	10.5471
CIT	20.1	179.19	232.17	200	0.5608572	7.000	1400	3.9260	9.8150
VOLUMES USED FOR SAMPLE PREPARATION FOR AMINO ACID ANALYSIS									
FOR URINE									
Sample:	10.0µL								
Isotope:	400µL								
FOR SERUM/PLASMA									
Sample:	10.0µL								
Isotope	400 µl								
FOR BLOOD-CARD									
Punch 1 x 1.495 µL sample = 1.495 µL (1 x 3.175 mm Punch)									
Use FC for preparation and calculation									

A3.5 Gradient of the mobile phases for the LC-QQQ analysis

LC-QQQ analysis was performed on 1200 series Agilent HPLC system linked to an Agilent 6410 MS/MS system. A C18 Zorbax SB-Aq column (150 mm × 2.1 mm × 3.5 µm) from Agilent was used for chromatographic separation of butylated compounds and kept at 30 °C throughout the analysis. Samples in the autosampler were kept at 4 °C. One µL sample was injected followed by the initiation of the programmed solvent gradient as discussed in S1 add shown in Table A3-1. The gradient for mobile phases A (0.1 % formic acid in water) and B (0.1 % formic acid in acetonitrile) is shown in Table A3-1.

Table A3–6: Mobile phase gradient for C18 column

Time (min)	Mobile Phase A (%)	Mobile Phase B (%)	Flow (mL/min)	rate
0	95	5		0.2
1	95	5		0.2
5	82	18		0.2
8	82	18		0.2
15	0	100		0.3
20	0	100		0.3
23	95	5		0.2

Each run were 23 minutes in length followed by a 7 minute post-run to re-equilibrate the column. The gas temperature for the ESI was kept at 300 °C at a flow rate of 7.5 L/minute. The capillary voltage was 3500 V and the nebulizer pressure remained 30 psi. The MRM transitions used to detect the target amino acids and the outcome of the analysis is shown in Table A3-7.

Table A3–7: Amino acids analysed in MRM mode with the following transitions using LC-MS/MS

Compound Name	Precursor Ion (m/z)	Product Ion (m/z)	Dwell time (s)	Fragmentor voltage (V)	Collision energy (V)
Citrulline IS	236.2	74.1	45	89	32
Arginine IS	235.2	74.3	45	103	32
Citrulline	232.2	70.1	45	89	32
Arginine	231.2	70.1	45	103	32
Histidine	212.1	110.1	45	89	16
Lysine IS	207.2	88.1	45	89	20
Lysine	203.2	84.1	45	89	20
Glutamine	203.1	84.1	45	89	20
Homocysteine	192.1	90.1	45	161	12
Ornithine	189.2	70.2	45	69	20
Asparagine	189.1	144.1	45	89	8
Creatine	188	90.1	45	89	16
Pyroglutamic acid	186	84.1	45	103	16
Cysteine	178.1	76	45	185	16
Threonine	176	74.2	45	84	12
Proline	172.1	70.2	45	94	20
Serine	162	60.1	45	79	12
GABA	160.1	87.1	45	79	8
Alanine	146.1	90.1	45	79	4
Glycine IS	134.1	78.2	45	65	4
Glycine	132.1	76.1	45	65	4
Homocystine	381.2	192.1	45	98	8
Cystine	353	130	45	118	16
Cystathionine	335.2	190.1	45	97	13
Glutamic acid IS	265	89.1	45	89	24
Tryptophan	261.2	159.1	45	94	16
Glutamic acid	260.2	84.1	45	89	24
Aspartic acid	246.2	144.1	45	98	8
Tyrosine	238.2	136.1	45	94	12
Phenylalanine IS	227.2	125.1	45	108	16
Leucine/Isoleucine	188.2	86.2	45	89	8
Valine IS	182.2	80.2	45	89	12
Valine	174.2	72.2	45	89	12

A3.6 Data processing and statistical analysis Cohort-1

Preprocessing was performed at a univariate level using log transformation and auto-scaling to place all variables on an equal footing (van den Berg *et al.* 2006). Outliers were identified based on the transformed and scaled data, using a 95% confidence region in a Hotelling's T^2 analysis in conjunction with a principal components analysis

(PCA) score plot (van den Berg *et al.* 2006; Eriksson *et al.* 2006), and confirmed as outliers based on clinical and analytical information. Univariate analysis of the two groups (without outliers) included fold changes (FC), Mann–Whitney (MW) *p*-values and effect sizes (ES). These statistics were calculated on the untransformed unscaled data (Cui and Churchill 2003).

Finally, multivariate analysis was based on the transformed and scaled data, excluding outliers. Unsupervised PCA was used to identify natural separation between groups, based on the first three components. In addition, 90% confidence ellipses were calculated for each group. Partial least squares discriminant analysis (PLS-DA) was the only supervised multivariate analysis performed on the scaled data. Scores for variables important in discrimination (VIP) were used in combination with univariate statistics to select a short list of significant variables for quantification and biological interpretation.

A3.7 Sample preparation and UPLS-MS/MS analysis

A3.7.1 Biogenic amine profiling

Sample collection, protein precipitation and further preparation of the samples up to storage was done exactly as for cohort-1. The amine platform used for the UPLS-MS/MS analysis covers amino acids and biogenic amines employing an Accq-tag derivatization strategy adapted from the protocol supplied by Waters. 5 µL of each serum sample was spiked with an internal standard solution, thiol amines are released from proteins and converted to reduced using TCEP. Then proteins are precipitated by the addition of methanol. The supernatant was transferred to a new eppendorf tube and taken to dryness in a speedvac. The residue was reconstituted in borate buffer (pH 8.5) with 6-Aminoquinolyl-N-hydroxysuccinim (AQC) reagent. After reaction, the vials were transferred to an autosampler tray and cooled to 10°C until the injection. 1.0 µL of the reaction mixture was injected into the UPLC-MS/MS system.

A3.7.2 UPLC-MS/MS Analyses and data processing

An ACQUITY UPLC system with autosampler (Waters, Etten-Leur, The Netherlands) was coupled online with a Xevo Tandem quadrupole mass spectrometer (Waters) operated using QuanLynx data acquisition software (version 4.1; Waters). The samples were analyzed by UPLC-MS/MS using an Accq-Tag Ultra column (Waters).

The Xevo TQ was used in the positive-ion electrospray mode and all analytes were monitored in Multiple Reaction Monitoring (MRM) using nominal mass resolution.

After LC-MS/MS analyses, peak integration was done using the provided TargetLynx software (Waters), by integration of assigned MRM peaks and the normalised relative ratios between metabolites and their corresponding internal standards were determined. For analysis of amino acids their ^{13}C ^{15}N -labeled analogs were used. For other amines, the closest-eluting internal standard was employed. Blank samples were used to correct for background, and in-house developed algorithms were applied using the pooled QC samples to compensate for shifts in the sensitivity of the mass spectrometer over the batch. Using the QC samples and duplicate samples, a double-quality-control approach was applied to include metabolites that were reliably measured by the individual metabolomics platforms by reporting and using only those metabolites for which both duplicate samples and QC samples showed an RSD < 30%.

A3.8 Outlier detection Cohort-2

Outliers for exposed and unexposed neonates were detected using equidistant binning data to identify cases with undue influences on the predictor models. This was done by the Hotelling's T^2 analysis (Figure A3-1a and A3-1c) and by assessing the principal components analysis (PCA) score plots for each of the experimental groups (Figure A3-1b and A3-1d) based on the same experimental data. Cases presenting outside of the confidence interval boundary (using a 90% confidence region) are classified as outliers. if detected both through the Hotelling's T^2 and PCA scores. However, no outliers were detected in any of the four groups, as illustrated in Figure A3-1 for the HIV-ART exposed neonates from cohort-1 and cohort-2.

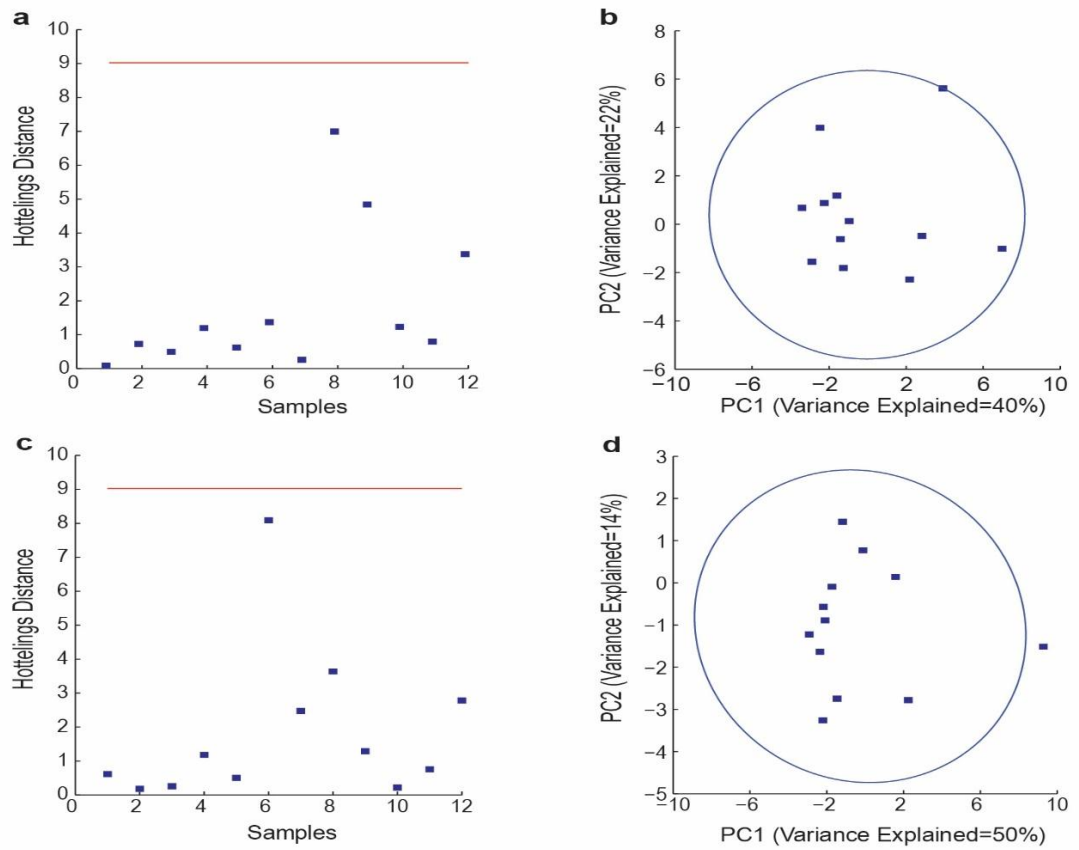


Figure A3-1: Outlier detection by means of a Hotelling's T^2 distance plot (**a** and **c**) and PCA scores plot (**b** and **d**) for the exposed cases from cohort-1 (**a** and **b**) and cohort-2 (**c** and **d**). Cases appearing both above the red line (Hotelling's T^2 plot) and outside the blue confidence interval boundary (PCA scores plot) would have been earmarked as outliers.

REFERENCES

Moutloatse, G. P., Bunders, M. J., van Reenen, M., Mason, S., Kuijpers, T. W., Engelke, U. F H. & Reinecke, C. J. (2016). Metabolic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood. *Metabolomics*, 12:175–189.

van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, K. A., & van der Werf, M. J. (2006). Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, 7(142), 1–15.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). Multi-and megavariable data analysis part I: basic principles and applications. Umetrics Inc.

Cui, X., & Churchill, G. A. (2003). Statistical tests for differential expression in cDNA microarray experiments. *Genome Biology*, 4(210), 1–10.

ANNEXURE 4: COPYRIGHT LICENCING AGREEMENTS

CHAPTER 2: IN UTERO EXPOSURE TO HIV-ART – A LITERATURE REVIEW

Figure 2–2: WHO: Adults and children estimated to be living with HIV in 2015

Licensee	Gontse Panache Moutloatse
Identity Number	214760
Licensed Date	29 November 2016
Excess	Website URL where WHO material is published : http://gamapserver.who.int/mapLibrary/Files/Maps/HIV_all_2015.png
Type of reuse	Reuse in a thesis/dissertation

Attached is the response to Me, Miss GP. Moutloatse from the permissions WHO international

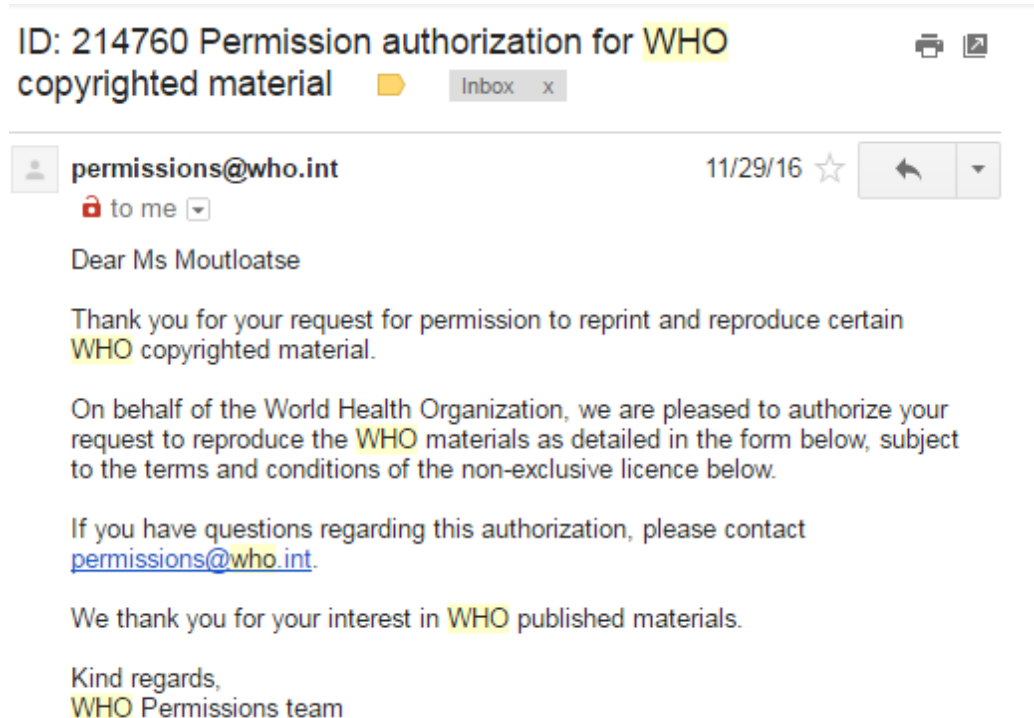


Figure 2–3: A conceptual representation of the effect of HIV and ART on mitochondrial metabolism

Licensee Gontse Panache Moutloatse

License Number 3997820178375

Licensed Date Nov 28, 2016

Licensed Publisher Elsevier

Licensed Title Metabolic and mitochondrial effects of antiretroviral drug exposure in pregnancy and postpartum: Implications for fetal and future health

Permission category Reuse in a thesis/dissertation

Order details and publisher terms and conditions are available by clicking on the link below:

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=e63b06ab-ffea-4b1b-97b2-c36e683b19bb>

CHAPTER 3: METABOLOMICS METHODOLOGY

Figure 3–4: Scientific advance may be seen as an interactive cycle linking knowledge and observations.

Licensee Gontse Panache Moutloatse

License Number 3998250346621

Licensed Date Nov 29, 2016

Licensed Publisher BioEssays

Licensed Title Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era

Permission category Reuse in a thesis/dissertation

Order details and publisher terms and conditions are available by clicking on the link below

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=3ca1cf47-7400-4005-a484-12053197400d>

CHAPTER 4: METABOLIC RISKS AT BIRTH FOLLOWING HIV-ART EXPOSURE: INSIGHTS FROM AN UNTARGETED METABOLOMICS STUDY

Licensee	Gontse P. Moutloatse
License Number	4020910670850
Licensed Date	Jan 02, 2017
Licensed Publication	Metabolomics
Licensed Title	Metabolic risks at birth of neonates exposed in utero to HIV-antiretroviral therapy relative to unexposed neonates: an NMR metabolomics study of cord blood
Type to use	Thesis/Dissertation

Order details and publisher terms and conditions are available by clicking on the link below

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=3370426e-a6d0-4438-956f-7ab7b072347d>

CHAPTER 5: REPEATABILITY IN THE GENERATION OF A MASS-SPECTROMETRIC-BASED METABOLOMICS DATA MATRIX

Figure 5–1: Scientific advance may be seen as an interactive cycle linking knowledge and observations

Licensee	Gontse Panache Moutloatse
----------	---------------------------

License Number :	4000800062645
------------------	---------------

Licensed Date Dec 02, 2016

Licensed Publisher Analyst

Licensed Title Data processing and the creation of a metabolomics data matrix

Permission category Reuse in a thesis/dissertation

Order details and publisher terms and conditions are available by clicking on the link below

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=3ca1cf47-7400-4005-a484-12053197400d>

Manuscript: **KEMREP: A New Qualitative Method for the Assessment of an Analyst's Ability to Generate a Metabolomics Data Matrix by Gas Chromatography– Mass Spectrometry**

Licensee Gontse P. Moutloatse

License Number 4063570858420

Licensed Date Mar 07, 2017

Licensed Publication Current Metabolomics

Licensed Title KEMREP: A New Qualitative Method for the Assessment of an Analyst's Ability to Generate a Metabolomics Data Matrix by Gas Chromatography– Mass Spectrometry

Type to use Thesis/Dissertation

Order details and publisher terms and conditions are available by clicking on the link below

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=00cc0a1e-0ab1-4785-b4bd-7b2b63c221e8>

CHAPTER 6: REPEATABILITY IN THE GENERATION OF A MASS-SPECTROMETRIC-BASED METABOLOMICS DATA MATRIX

Manuscript: Metabolic risks of neonates at birth following in utero exposure to HIV-ART: the amino acid profile of cord blood

Licensee Gontse P. Moutloatse

License Number 4174711111246

Licensed Date Aug 23, 2017

Licensed Publication Metabolomics

Licensed Title Metabolic risks of neonates at birth following in utero exposure to HIV-ART: the amino acid profile of cord blood

Type to use Thesis/Dissertation

Order details and publisher terms and conditions are available by clicking on the link below

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=ebc10864-a133-4324-abae-ffb98687e6ca>

Manuscript: Fetal metabolic stress disrupts immune homeostasis and induces pro-inflammatory responses in HIV and cART-exposed infants

Licensee Gontse P. Moutloatse

License Number 4145290944017

Licensed Date Jul 10, 2017

Licensed Publication Journal of Infectious Diseases

Licensed Title Fetal metabolic stress disrupts immune homeostasis and induces pro-inflammatory responses in HIV-1 and cART-exposed infants

Type to use Thesis/Dissertation

Order details and publisher terms and conditions are available by clicking on the link below

<https://s100.copyright.com/MyAccount/viewLicenseDetails?ref=555f16ca-a036-4284-9a00-920404d0751e>

CHAPTER 7: REPEATABILITY IN THE GENERATION OF A MASS-SPECTROMETRIC-BASED METABOLOMICS DATA MATRIX

Manuscript: 1H-NMR Metabolomics Studies on Neonatal Cord Blood

Submitted: *Current Metabolomics*, 2016